DOCUMENT RESUME

ED 443 610 RC 022 302

TITLE	Assessment Monograph: A Special Report to the Rural School and Community Trust.
INSTITUTION	Harvard Univ., Cambridge, MA. Graduate School of Education.
SPONS AGENCY	Rural School and Community Trust, Washington, DC.
PUB DATE	1999-10-00
NOTE	83p.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	EDRS Price MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	*Alternative Assessment; Educational Practices; Elementary Secondary Education; Evaluation Criteria; *Evaluation Methods; *Performance Based Assessment; Rural Schools; School Community Relationship; *Scoring Rubrics; Standardized Tests; *Student Evaluation; Test Reliability; Test Validity

ABSTRACT

This guide to different/alternative assessment was prepared for participants in the Rural School and Community Trust, formerly the Annenberg Rural Challenge. Given that Rural Trust participants focus on local place-based education and the relationship between rural schools and their communities, they are not, by and large, enamored of the externally-developed testing that has come to dominate schooling in America. This monograph briefly explains what's wrong with standardized tests and presents a detailed exposition of several assessment practices that aim to demonstrate student understanding through real-world problem solving and products of student work. These different/alternative assessment practices are also known as authentic assessments, performance assessments, educative assessments (formative evaluation), exhibitions, and portfolio assessment, among others. The monograph outlines principles of alternative assessments, recommends practices for designing different/alternative assessments, offers suggestions for constructing rubrics, and discusses ways to engage students in the evaluation criteria for quality performance. Specific examples are provided from various rural school communities that are constructing assessment procedures to reflect the quality of student work while centering student learning on the local community. An appendix presents a sample of a fifth-grade writing portfolio in a Vermont school. (Contains 16 references and 3 related Web sites.) (SV)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 443 610

Assessment Monograph

A Special Report to the Rural School and Community Trust

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
Carla Fontaine
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
2B

BEST COPY AVAILABLE

Research and Evaluation Program
Harvard University Graduate School of Education

RC 022302

Assessment Monograph

A Special Report to the Rural School and Community Trust

October 1999

Dedicated to the many teachers,
students, parents and community people
who are engaged in the work of
the Rural Trust

Printed on recycled paper

TABLE OF CONTENTS

Preface	1
Introduction	3
Different/Alternative Assessments	4
What's Wrong with the Tests?	5
New (Old) Alternatives	6
Sample Performance Tasks, Exhibitions, Portfolio Assessments	7
Principles of Different/Alternative Assessment	9
Assessing the Validity and Reliability of Tests and Different/Alternative Assessments	10
Designing Different/Alternative Assessment	12
Examples of Practices in Designing Different/ Alternative Assessment Tasks	14
The Molly Ockett Middle School (Maine) Story	14
Recommended Practices for Designing Different/ Alternative Assessments	18
Articulated Goals of Instruction	18
Valid and Appropriate Evidence	20
Criteria or Standards for Assessment	22
Student Work in Writing and Mathematics	22
Designing Assessment Tasks	23
A Project in Cost-effective Design	23
8th Grade Mathematics Performance Task	25
Predicting Changes	26
Part of a Civil Rights Unit	27
9th Grade Science Energy Unit	28
Rubrics for Scoring Product/s	29
Heritage High School Senior Essay	30
Writing Assessment Criteria	31
Six Traits for Teaching/Assessing Writing	32
Six Trait Analytical Scoring Guide	33

Three Trait Writing Rubric 35
Mathematics Problem Solution Generalized Holistic Rubric 37
Analytic Trait Rubric for 5th Grade Science Experiments 38
Heroes Exhibition Rubric 39
Task: Research Paper 40
Rubric for Assessing a Speech 41

Suggestions for Constructing Rubrics 42
Designing Your Own Rubric 42
Central Park East Secondary School Evaluation Guide 46

Engaging Students in Evaluation Criteria for Quality Performances 47
Self-Evaluation Check List 48
Project Work Reflections 48
The Trial of Simeone Bushnell 49
Evaluation of the Trial of Simeone Bushnell 50

Preparing Evaluators 51

Portfolios as Assessment Tools 51
Vermont Language Arts Portfolio Assessment 53

Case Study of Implementing Different/Alternative Assessments 54

How Can We Observe Student Learning with Performance Assessments? 61

Public Accounting 66

Closing Statement 66

Bibliography 67

Appendix (5th Grade Writing Portfolio) 69

Rural School and Community Trust Assessment Monograph[1]

Preface

This "guide" has been prepared specifically for Rural Trust participants, though it may be useful to other educators as well. Participants in the Rural Trust are not, by and large, enamored with all the externally-developed testing that has come to dominate schooling in America. Given their enlarging commitments to educational practices that give particular attention to what is local, directed at both school and community well-being, toward mutuality, the work being done is growing increasingly distant from the various content specifications embedded in the external testing being mandated by county and state bureaucracies.

The belief is that the local focus makes what is studied more connected to the lives of students, families and communities and brings to students greater intellectual engagement while providing them higher levels of skills relating to the traditional subject matters of reading, writing, math and science as well as critical habits of mind that relate to genuine citizenship -- thinking, decision-making, planning, presenting and valuing of local culture. Rural Trust participants believe these skills and habits of mind are more fully and authentically displayed through long-term projects, in work that has a public purpose and is made available through public forums, than through tests which are most often unrelated to the larger purposes of their schools.

This guide provides a short critique of these tests (norm-referenced and criterion-referenced measures) along with some related definitions of testing terms. This particular discussion is meant to enlarge understanding, to provide Rural Trust participants with more information to buttress their more intuitive knowledge about the inadequacy, even unsalutary effects of these tests.

It also serves to acknowledge that those in Rural Trust schools must argue for, and make use of, assessment practices that are more potent, that support their students and the teaching-learning exchange more productively, that inform their local communities more fully. In the best of worlds, teachers and schools would care almost exclusively about their own children, parents and local community members. These are the people who need to understand fully what is being taught and what is being learned, understood, cared about. This is where accountability should truly matter. It is at this local level that student work in all its dimensions can be made most visible, fully public. This is not to suggest, of course, that others beyond the local setting do not have legitimate accountability interests but accountability must begin locally to have any genuine credibility.

We provide in this guide ways in which teachers and schools have begun to make assessment more systematic, using the results to report to their communities as well as to governmental agencies beyond their communities. We focus heavily on assessment that is rooted in actual work over time, mostly around portfolios and various kinds of rubrics for evaluating work. The stress is on situated work over time, the kind of work that is central to ongoing teaching-learning practices rather than one-shot tests that are unsituated and unrelated to the ongoing instructional programs of schools, and to which students are seldom committed.

The guide is not intended to define what schools should necessarily do. In the end, teachers in particular schools, working with students, parents and community people, will have to develop their assessment efforts in a manner that makes sense to them, that meets their own well defined purposes. The hope is that the discussion and the various examples provided in this guide will be useful in support of that purpose.

This monograph had its origin in a request from Rural Trust Project Directors who wished to provide teachers and schools with updated information about assessment practices that might be more accommodating to Rural Trust-oriented work than the various tests. In the course of completing the

---

[1] As of October 1, 1999, the Rural School and Community Trust is the new name for what was formerly the Annenberg Rural Challenge.

BEST COPY AVAILABLE

document, many Project Directors provided constructive counsel and critique. We thank, in particular, David Ruff and Edd Diden. The document is better for their interest and suggestions.

I need especially to acknowledge Polly Ulichny who carried the burden of researching and writing much of this assessment document. Her knowledge of schools and assessment practices made a great difference. Carla Fontaine contributed to the text through her work with several teachers in Rural Trust sites. She also served in a critical editorial role, making the text more accessible to the wide Rural Trust audience we wish to reach. Several Rural Trust teachers contributed to the document by sharing materials and the stories of their assessment practices for use in the monograph. They are as follows: Tom McLaughlin, Pat Higgins and Bob Cobb of the Molly Ockett Middle School in Fryeburg, Maine; Liza Finkel at Noble High School in Berwick, Maine; Mitch Mendosa from Anderson Valley, California; Judy Cross and Regina Headden at the Wartburg Central High School (Tennessee); and Margaret MacLean and Kathy Renfrew of the Peacham School (Vermont). I wish also to acknowledge Jan Still who prepared much of the text.

We trust the monograph will be useful to Rural Trust participants who continually inspire us with the quality of their work and their vital commitments to children, young people and communities.

Vito Perrone
Cambridge, Massachusetts
October 1999

We pass through this world but once. Few tragedies can be more extensive than the stunting of life, few injustices deeper than the denial of an opportunity to strive or even to hope, by a limit imposed from without but falsely identified as lying within.

Stephen Jay Gould, The Mismeasure of Man

Introduction

Assessment plays a powerful role in education. At its best, it informs, guides and supports the growth of students as it provides teachers with critical directions for their ongoing teaching. What teachers and schools choose to assess determines to a great extent not only what gets taught in schools but how it gets taught. Students have most commonly experienced assessment through teacher-made tests, quizzes, district or textbook unit tests, completion of projects and evidence of classroom participation and various standardized tests.[2] The purposes of these assessments have varied but, fundamentally, they have been utilized to keep track of student learning. They have also been used, in many settings, to support the various sorting and selection decisions made by those in schools. For the latter purposes, scores from norm and criterion-referenced tests have come to matter the most. Increasingly, however, we are seeing these standardized tests assume greater power -- being used more frequently for the purpose of auditing the performance of an entire school or school district and determining for students the successful or unsuccessful completion of mandatory schooling (what has come to be called the new high stakes quality of testing).

In this historical, cultural moment in which market values, accountability and bottom-line thinking have come to dominate the public discourse, the stakes are rising dramatically and are increasingly attached to student measures that have the appearance of objectivity and a lack of bias, and produce standardized results. This increased significance accorded to standardized tests has become a major concern to many Rural Trust educators, parents and students who have come to see the external testing as a growing impediment to the more productive place and performance-oriented work to which they are committed.[3] Their objections are grounded in a critique of externally developed norm and criterion-referenced tests being used as a major determinant of educational practices and curriculum. They argue that a climate is being created to support teaching almost exclusively to the tests, thereby curtailing local determination of curriculum. They also argue that standardized tests are neither as objective, valid or bias free as some would believe.[4] Moreover, they question whether the kinds of

---

[2] We include standardized tests in this outline of what students experience because students have, over the past 25 years, had to take an increasing array of such tests. While they haven't always had a direct influence on grades, promotion or graduation, the testing landscape is changing. In this monograph, we will discuss the changes.

[3] In suggesting that the tests are an impediment, we do not wish to say that some alignment of place-oriented work and various state curriculum standards and tests cannot be accomplished. To work at such alignment, however, seems ultimately wasteful. The discourse ought to be about enlarging the possibilities of addressing local concerns, making larger school-community connections, working toward a larger mutuality. We need, in this regard, to keep asking: "Will our assessment directions enlarge the educational possibilities for our students?" By and large, the tests limit the possibilities.

[4] Given the terms that surround various standardized tests -- objectivity, standardized, reliability and validity -- and the aura of science, it might be useful to provide their meanings in basic, non-technical terms.

A test is considered *objective* if everyone takes it under the same basic conditions. The multiple-choice format, buttressed by a single "right" answer pattern, supports objectivity. But objectivity has nothing to do with whether a test is fair, contains items of importance, or has ambiguous questions and answers. Objectivity has, in other words, *no relationship to quality*.

A test is *standardized* if norms have been established. Whether the norm populations are representative in more than a statistical sense is not the defining characteristic. This term, as in the case for objectivity, *has no relationship to quality*.

Reliability relates to the consistency of the test: How close are the results for an individual or group at two different testings? Or how close are the scores of individuals on two different forms of a particular test? Reliability is rather simple to establish. But a test can have very high reliability (most popular standardized achievement tests

answers that can be assessed unequivocally by a simple scoring system (conducted primarily by machines) should set the standard for what comes to be valued as an end of education. Grant Wiggins, a proponent of alternative assessment measures, claims that "[w]e sacrifice our aims [as educators] and our children's intellectual needs when we test what is easy to test rather than the complex and rich tasks that we value in our classrooms and that are at the heart of our curriculum."[5]

Rural Trust communities have been asking schools to prepare students to be genuine citizens of their places, wise consumers and producers of knowledge, practiced in the complex tasks that are required of adults in their communities, in the workplace, and in the ever more accessible global village. Such ends are not tapped by the various tests being used by states. Much of the testing now emerging appears, in fact, to work against the kinds of understandings and competencies that Rural Trust schools and communities hope to cultivate in students. To support the larger ends Rural Trust schools and communities are working toward will require compatible assessment practices, in which students actually do something, construct answers, perform critical tasks, produce something that has meaning, that is fully embedded in the ongoing teaching-learning exchange.

We intend in this monograph to address more meaningful and valuable means of getting closer to students' understandings of what teachers, schools and communities most want them to learn as well as provide information to teachers and students that is critical to improving teaching and learning. In this respect, assessment practices must be understandable to the students, seen as valuable. Current tests don't match that expectation. We will present a detailed exposition of several assessment practices that attempt to capture student understanding that is deep, complex and demonstrated in products of student work rather than in paper and pencil tests. We will also provide some background to the ideas embodied in these different/alternative assessment practices, discuss the practicalities of undertaking some of these assessment practices, and provide specific examples from a variety of school communities, including various Rural Trust sites, that are engaged in constructing assessment procedures that reflect the quality of work students are producing as they center their learning on their local communities.

Different/Alternative Assessments

What we call in this monograph different/alternative assessment practices are also known by other names -- authentic assessments, performance assessments, educative assessments, exhibitions and portfolio assessments, among others. The variety of names refers to aspects of the process that are accentuated (as in the case of "performance"), the mirroring of learning tasks to real-life problem-solving practices ("authentic"), or to the array of student work that is to be assessed (an "exhibition" of

---

carry reliability coefficients of .87 to .93) and yet be a very poor test, measuring little considered important to large numbers of people who use or take the test. The latter point relates to *validity*, which has an interdependent relationship with reliability. Validity doesn't, however, receive as much attention.

Validity refers, at the most simple level, to the degree to which a test measures what it is supposed to measure and/or the degree to which the scores derived from a particular test are related to what the test is supposed to be measuring. In other words, what are the inferences that can be drawn from the test scores? Validity, unlike reliability, is difficult to establish with any authority. It is typically determined by having an expert (or experts) examine a particular test and provide the equivalent of an imprimatur. This is content validation by opinion. Content validation is the focus of most standardized tests used in elementary and secondary schools. (Those who suggest that the tests often contain bias are, in essence, questioning content validity by claiming that the test content is not representative of the socio-educational experience of particular populations.) Validity (face validity) is also established by comparing the test results with other measures, i.e., other tests, grades or teacher judgments.

[5] Grant Wiggins, Educative assessment: Designing assessments to inform and improve student performance. (San Francisco: Jossey-Bass, 1998), 7. Wiggins argues that most standardized tests don't get very close to what teachers and students believe matters, to what students actually value and understand. We should also note that scores on these standardized tests correlate most fully with the socio-economic background of the students. Those from wealthier communities tend to score much higher than students from poorer communities. This factor alone -- and it has been consistent over the entire history of standardized tests -- should cause us to see the tests as socio-culturally problematic.

understanding, a portfolio of work). While "alternative assessment" is the common expression of those who support performance and portfolio assessment practices, we have chosen to use *different* alongside *alternative*. To use alternative exclusively seems to give too much legitimacy to the processes currently dominating assessment in schools. Embedded in the philosophy of these different/alternative assessments is another understanding of teaching and learning.

What's Wrong with the Tests

A great deal has been written about the shortcomings of standardized tests as an adequate measure of what students know and think.[6] Some forms of standardized testing are criticized more than others, but all large scale standardized assessments are faulted for their excessive focus on what students *know about* a subject or process rather than a demonstration of *know how*, which requires applying knowledge appropriately in some real-life or realistic context. Standardized tests have traditionally required students to choose the "right" answer from a predetermined set of possible answers. This method can only tap a *know about* ability, which is heavily dependent on item memory, ease and speed of recall, and having been exposed to the specific item being sought. Even worse, because of the almost exclusive reliance on standardized assessments to sort, judge, and rank students, they can exert pressure on the entire educational system to keep instruction and learning at a superficial level, coverage becoming more important than uncoverage.[7] Surface level recognition, recall, and rote application of common procedures is privileged in many classrooms in order to cover the quantity of topics and items that a student might encounter during the one-shot, end-of-the-year test that is then whisked away to a central grading facility which disseminates the results several months later -- too late to offer any critical information to the students' current teachers.[8]

While few would argue that this kind of externally developed testing actually captures what students understand or know how to do in more complex real-life situations in which basic knowledge must be applied to complete an important task or solve a problem, it, nevertheless, prevails as the dominant form of assessment because of the particular qualities that exist with this type of testing. Multiple-choice questions, for example, can be easily scored in large numbers and student performance on the same test can be compared.[9] Data that can be easily quantifiable seem simpler to attend to than "messy and complex" data -- promoting the belief that if you can't count it, it doesn't count. A student's performance on these tests is evaluated on the basis of how it compares to all the other test-takers rather than against stable, articulated criteria that specify what quality performance on the test would be or even what is worth knowing. Furthermore, test-makers contend that secrecy of actual test items and how an individual scored on them is required in order to permit continual comparative evaluation of student scores as well as to safeguard the "reliability" of test results. This latter characteristic insures that the results of the tests will not inform or educate students or their teachers about what particular students got right or wrong or how they might improve performance. It merely tells them how they fared in the competitive rankings of student performance. It is not difficult, under these circumstances, to see why the

---

[6] We have chosen to put up-front some of the criticisms of standardized tests to provide Rural Challenge participants with some useful information, to make clear our ongoing educative purposes. We should note that throughout the history of standardized testing, criticism has been large. It is not a recent phenomenon.

[7] It is unfortunate that test practice and preparation have come to dominate so much of students' time in schools. Such a focus might produce higher test scores as it guarantees an even more limited education.

[8] We have obviously only touched the surface here. Ambiguities in tests are commonplace. How a child is feeling on the day the test is given can affect a score. Test familiarity matters. Mechanical errors are easy to make. Cultural and linguistic differences often influence children's understandings of various questions. Such factors mask greatly what students, in fact, understand. To make critical decisions on the basis of these tests is not educationally defensible. (We should acknowledge that most test-makers also argue against the use of a single test for educational decision-making.)

[9] Many of the state tests being developed include some open-ended items. They appear, however, to have much of the quality of multiple-choice items. Further, these tests usually include a writing task. The unsituated quality of such exercises, though, is not conducive to inspiring students' best work.

BEST COPY AVAILABLE

information acquired from these tests is not informative of what a particular student can do with his or her understandings nor does it help improve teaching and learning to enhance student performance.

Why, one might ask, does the educational establishment and the public put so much stock in these procedures, to the point that, across the country, scores on standardized tests are increasingly determining whether students pass or fail their grade or even whether they receive a regular high school diploma or merely a certificate of attendance? The answer lies in a historical, cultural preference for basing educational policy on technical solutions, simple numbers, so-called objective measures, and a belief that competition among individuals not only reveals excellence but actually produces it. This classic free market analogy is increasingly underlying most public policy decisions.

Teacher-made tests, of course, may suffer from some of the same inadequacies of standardized tests. Many, in fact, rely on easy to correct answers that tap the same kinds of superficial, coverage-oriented knowledge as standardized tests. The culture of testing is so dominated by the habit of auditing what is in a student's head at a given point in time, and of permitting a comparative judgment of student knowledge in a particular subject area, that teacher-made tests may vary little in format from standardized tests that require choosing the right answer. Further, even in settings in which teachers use formats other than simple recall, much of the assessment is rooted solely in teacher judgment.[10] In this regard, what counts as an error, the severity of types of errors, the day of the test, the topic, and so on, can all still influence the evaluation process. Since, in most schools, there is little opportunity or occasion to compare either the tests or the grades that students receive from different teachers, the value of teacher-made assessment as a basis for comparing the work of students across classrooms or schools or making school-to-school comparisons is almost non-existent, even though the results in a particular classroom are more likely related to what teachers actually teach (which we would argue is important).[11]

In this brief discussion we have tried to convey how the various constancies of assessment/testing have created a limiting view of the process: one designed to audit and rank students according to recall of superficial facts and ideas. Increasingly, however, educators have begun to advocate different forms of assessment that have potential for serving more meaningful purposes in teaching and learning.[12]

New (Old) Alternatives

Norman Frederiksen, a senior researcher at the Educational Testing Service (ETS) addressed some of what we need to do when he stated that "[m]ost of the important problems one faces in real life are ill-structured, as are all the really important social, political, and scientific problems in the world today. But ill-structured problems are not found in standardized achievement tests.... We need a much broader conception of what a test is if we are to use test information in improving educational outcomes."[13] Since what and how we choose to assess students' proficiency determines in large measure what and how we teach the students, basing our assessment practices on what we truly value in an education matters greatly. The long tradition of progressive educators who advocated more educative, informative, and nuanced assessment techniques has created a number of practices that go beyond, and are more productive than, standardized tests and the arbitrariness of too many teacher-made pen and paper tests. These practices include strategies for assessing student understanding that are made visible in public

---

[10] While we have more trust in teacher judgment about student understanding than in the results of standardized test scores, we also must acknowledge that teacher judgment alone has limitations when made a central element of an accountability system.

[11] We acknowledge that there are school settings in which teachers collectively discuss essays written in relation to various subject matters, in the process expanding everyone's understandings of what constitutes higher and lower quality.

[12] We do not wish to imply here that the different forms of assessment we explore in this monograph are necessarily new. Many progressive educators at the turn of the century were involved in close observation of students, documenting carefully their ongoing progress as learners by maintaining work folders (what are now described as portfolios). And performance activities (what are now discussed as exhibitions) were fairly commonplace.

[13] N. Frederiksen, "The Real Test Bias, "*American Psychologist* 39 (1984) as quoted in Grant Wiggins, Assessing student performance: Exploring the purpose and limits of testing (San Francisco: Jossey-Bass, 1993), p. 4.

exhibitions of expertise, projects, laboratory experiments, portfolios of work and ongoing records of various learning activities.[14] What do some of these assessment tasks look like?

While student performance is likely best observed/documented within ongoing instructional activities, some schools have begun to use well-constructed performance assessment tasks to get closer to students' ability to use knowledge, to demonstrate that they can go beyond what has been presented in their texts or classes. These kinds of performance tasks are often used as weekly tests or unit tests were used in the past. They are often used, as well, in schools that have established portfolio requirements as demonstrations of progress or fuller understanding.

SAMPLE PERFORMANCE TASKS, EXHIBITIONS, PORTFOLIO ASSESSMENTS

- *Towering Trees*

How can the theater tech crew backstage best meet scenery needs of the community theater group? Students must determine a linear function relating height and another appropriate variable to construct trees for a stage production. Using that equation, the students must select the most appropriate and cost-effective size of plywood to construct props for the Theater Arts Guild productions. The students must support their findings in an oral presentation and written report to the Guild using apt graphics and other support material.

- *Troubled Waters – Headline News*

"CONFLICT ON THE COAST -- FRIENDLY RIVER BECOMES DEADLY
Major fish kills....100,000 fish wash ashore; child returns from swim with hives; seafood harvest plummets; and scientist discovers fish-killing invader!" What is the problem? In this task, students take on the role of the "Riverkeeper," whose job is to maintain healthy rivers. Students will analyze data gathered along the Anywhere River, pinpoint pollution hotspots, and outline research to determine potential sources. Two letters will be drafted: (1) an informative letter to fishermen, and (2) a letter to a governmental agency, convincing them to address the situation.[15]

- *Heroes Exhibition*

What kind of heroes would you prescribe for our society?
After whom should we pattern our lives?

For this exhibition, you will create and justify a collection of heroes for our time and culture. In your project you will define your idea of a hero for late 20th century America, based on your own interpretation and synthesis of class activities. You will select a group of persons -- mythic or real, living or dead, famous or known only to a few -- who exemplify that definition. You will have the opportunity to choose the format of your project from a list which follows. Finally, you will make an oral presentation of your collection, first to a small group of your peers and then to the whole class, during which you will explain and defend your ideas more completely, focusing on our central question: Who should our heroes be?

---

[14] Organizations that have developed and promoted these alternatives in recent years include the Coalition of Essential Schools, the Prospect School, the North Dakota Study Group, Project Zero, the Foxfire Outreach Network, several professional associations, and other networks of progressive schools.

[15] These two performance tasks were developed by the Center on Learning, Assessment, and School Structure (CLASS) and appear in Wiggins, *Educative Assessment*, p. 143.

[In relation to the task]
Decide on a format for your project. You may choose from the following list:

a.	*Anthology of stories and poems, original and found*	*[includes author's introduction]*
b.	*Series of monologues*	*[includes performer's introduction]*
c.	*Expository speech*	*[includes intro, body, conclusion]*
d.	*Illustrated calendar*	*[plus promotional copy]*
e.	*Set of trading cards*	*[plus promotional copy]*
f.	*Illustrated magazine or comic book*	*[includes author's introduction]*
g.	*Hypercard stack*	*[includes author's introduction]*
h.	*Mural*	*[with legend and artist's notes]*

[And]
Write an exhibition proposal, which includes:
a. a bibliography of sources you have consulted and plan to consult
b. a working list of heroes you plan to include
NOTE: Both 'a' and 'b' may change as you develop your exhibition
c. the format of your project
d. a list of materials needed
e. a timeline, including estimated dates of presentation to your work group and to the class. [16]

- *Autobiographic Portfolio/Semester I*

So that our seniors understand their skills, strengths and talents better, and in order to enhance the college admissions process, they will create and compile an autobiographic portfolio, to include written and visual components.

Essential questions, like the ones below, will inspire students to begin thinking about the scope and nature of their portfolios. Students may write their own questions which address issues of identity, place, society, family, etc. These questions should be used as a "starting place" for their autobiographical writings.

1. How does the geographical region(s) in which you are/were raised help define who you are?

---

[16] This particular exhibition was developed for the Humanities curriculum at Fulton Valley Prep at Piner High School. A fuller treatment is presented in Bill Johnson, The performance assessment handbook: Designs from the field and guidelines for the territory ahead, Vol. 2 (Princeton, NJ: Eye on Education, 1996), 86-95.

2. *How do I want to be remembered?*
3. *How do I fit into a pluralistic society?*
4. *How has my relationship to my neighborhood, church, family, school, or peers shaped who I am?*

Assessment of the portfolio and the Senior Thesis will be by a committee consisting of the following people: one outside adult, one faculty advisory (assigned), one junior from your advisor group, a second faculty member, and an outside "expert." (EC = the student's Exhibition Committee)

The portfolio will include:

1. *College Essay(s) ....*
2. *A Post-Graduate Plan which accounts for the whole person (Body, Mind, Spirit)....*
3. *A Written Autobiography....*
4. *A Non-Written Component [an artistic or technological medium expression of some aspect of self]*
5. *The Proposal for Senior Thesis*
6. *Written Reflection on the question, "How does what I do reflect who I am?"*
7. *Rubric for Assessment of Autobiographic Portfolio*[17]

Principles of Different/Alternative Assessments

A number of principles underscore the philosophy of different/alternative assessments and are often reflected in the various names given to these assessment practices. Some of the most notable are the following:

1. *What we assess must mirror what we consider as valuable outcomes of the educational process.* If we are committed to preparing students in habits of mind that support thoughtful, democratic participation in one's community, wider society and the world, we must engage them in tasks that support such values and assess their ability to engage in them.
2. *Assessment tasks should resemble real-life problem-solving,* requiring the application of knowledge to fit a purpose, deciding how to tackle a problem, experimentation, demonstrating knowledge of how to accomplish a goal. This principle is sometimes captured by the label "authentic" assessment measures, although that term has drawn criticism since assessment tasks are often constructed tasks, not "real" or authentic instances of problem-solving in life. A more fitting term may be "direct" assessment, meaning that the complex task of problem solving as it occurs in the real world is simulated in an assessment task and students are assessed on their ability to perform. This is in contrast to indirect assessments of complex behaviors by means of easy-to-test skills, such as knowledge of word meaning to assess reading, simple calculations to assess mathematical understanding, selecting or reporting facts and interpretations that can be put into questions with only one right answer to assess understanding. No matter what they are called, however, authentic or direct assessments of complex tasks imply a messier, multi-layered effort that will result in a product, not necessarily a single right answer. The evaluation of that product must, therefore, be as multi-layered and complex as the task warrants. (We will discuss this in detail below.)
3. *Trustworthiness or validity of the evaluation of performance assessments depends on practices and procedures* that include specifying and making public the criteria for quality performances, relying on multiple sources of student-generated data, including longitudinal evidence to evaluate

[17] The full version of this Senior Portfolio/Exhibition Handbook was developed by staff at Heathwood Hall Episcopal School and is presented in Johnson, *The Performance Assessment Handbook Vol. 2,* 104 -113.

BEST COPY AVAILABLE

learning, ability and know how, and using multiple evaluators practiced in the methods of assessment demanded by complex performances.[18]

4. *Assessment information can be used to improve the teaching and learning* that occurs in a particular classroom as well as get close to a student's understandings at a given point in time. The feedback from self-evaluation that comes from students knowing and applying the criteria for quality performance of a task, peer evaluation from fellow students, and teachers and outside evaluators' ratings, using the same criteria, informs the student not only about how well they are progressing, but more importantly, what they can do to improve performance. An examination of complex tasks and performances also informs teachers of next steps or adaptations in instruction. This formative or educative function has the potential of altering the culture of assessment from an auditing and ranking practice to one of informing instruction and learning.
5. *The audience for whom the various assessment tasks are designed should be as authentic as possible.* While this element is not always feasible, the work of the Rural Trust holds this principle as central to its mission. As school and community become more interdependent and collectively address the issues of sustainability and quality of life in the local area, more of the assessment performances and exhibitions should have as their audience the community, the people affected by many of the academic projects in which students are involved.

Assessing the Validity and Reliability of Tests and Different/Alternative Assessments

Psychometricians, whose work "tests the tests," have established criteria for determining the validity and reliability of specific tests so that they may be administered widely (nationally and internationally) and generate credible results.[19] Currently, as we noted earlier, we are seeing the implementation of high stakes tests at state and district levels that determine whether a student is promoted or not, and even whether schools are permitted to function semi-autonomously or be taken over by a regulatory body to improve test scores. To support these various decisions based on test scores, standards of objectivity, reliability, and validity are typically invoked. What enables a test to meet some form of these standards and be efficient and economical to administer to large numbers of students essentially creates our current multiple-choice standardized tests. Different/alternative assessments are unlikely to supplant these tests as large scale assessments because they cannot easily be standardized or administered as efficiently and economically as mass-produced, machine-scored standardized tests.[20] Yet, they provide more detailed information about student learning than standardized tests allow. This, of course, should matter. We discuss below these various test-related criteria -- objectivity, reliability and validity -- from the perspective of both kinds of assessments to see how each responds to their demands.[21]

Objectivity is often defined as "a lack of bias in tests." Since all test-takers receive the same test, the same amount of time to complete it, and the scoring is done mechanically, with no intervention of human judgment, standardized tests claim a lack of subjective bias. Gertrude Conlan notes, however, that "[w]ith the term *objective* comes the connotation of scientific accuracy and precision.... What is forgotten in all these associations and misconceptions is that *objective* refers to the method of scoring, not

---

[18] We refer to evaluators at many points in this discussion. Evaluators can be teachers, students, parents, community members, specialists in the work being undertaken by students. For purposes of evaluation, beyond a formative stage, it is critical to have work examined by multiple evaluators.

[19] The technology of this large scale assessment work is relatively new, having its beginnings early in the 20th century. The science of psychometrics evolved to support the development of tests and to establish standards for their use. It should be noted again here that few testing professionals support the use of a test score as a single, even principal, determiner of promotion, graduation or placement into or out of various educational opportunities.

[20] While we don't see these different/alternative assessment practices supplanting these large-scale assessment activities anytime soon, we believe they can be the basis for schools seeking waivers from such systems and ultimately for schools to make a case for more appropriate modes of assessment on a larger scale. As a public policy effort, these current high stakes practices are not defensible educationally.

[21] While we provided earlier some basic definitions surrounding standardized tests, we extend the discussion here.

to the test itself."[22] The very fact that the test is the same for all students and administered under similar conditions is, in reality, a powerful source of bias -- that of the test-makers. Since students come from and possess a variety of backgrounds, local customs, knowledge and values, not to mention that their schooling varies in content and procedures, no one test can fairly assess what diverse students know and can do. This inherent lack of fairness is produced, in fact, because of the "objective" techniques of test production and administration.

Well designed different/alternative assessments result in student products that are unique accomplishments of individuals and groups of students. There is no "right answer" that can be checked off by any scorer. There are ways that different assessments can be evaluated, however, that adhere to a fairness principle. For example, in settings that regularly use exhibitions, products or portfolios, criteria for excellent, adequate and insufficient performance are articulated for specific traits that are to be evaluated. These are most often presented in the form of *rubrics* that are made public, that students can use to guide their own performance, and can provide, in the process, the framework for feedback on works in progress as well as final products. (The design of rubrics and exemplars will be explored in detail below). Teachers and external evaluators regularly discuss their ratings/evaluations in terms of the products and the rubrics. This conversation helps build a community of readers and evaluators of student work able to assess work according to specific rubrics. It also assures a general reliability of assessment across evaluators. Keeping track of how teachers and external evaluators rate the same work according to the same rubric serves as a tool for developing consistent application of scoring criteria. When discrepancies occur in the assignment of scores, those involved in the evaluation can discuss the rationale for their scoring to reach agreement on how the scores and rubrics are to be interpreted. It, therefore, encourages practices that anchor evaluations to specific criteria (via rubrics), thus ensuring inter-rater reliability through consensus-building evaluation dialogues.

Reliability relates to consistency across time and on related parts of a test and represents another criterion that tests are supposed to meet. This criterion basically asks if the results of a test are consistent with knowledge of the test taker. One way that reliability is statistically measured for standardized tests is by selecting items within the same test that ask for the same knowledge to see if individual test-takers respond consistently to the two or more items. Another is to look at test results of the same test over time across large samples of test takers to see if the scores follow a similar distribution pattern. The statistical procedures that demonstrate this kind of consistency of response patterns may say something about the test items, but they can never insure that a particular score of an individual test-taker is a reliable indicator of his or her knowledge.[23]

The question of reliability for different/alternative assessments rests on the practices of evaluation and on the design of the assessment tasks. As we mentioned earlier, the quality of the conversations, preparation of evaluators, and general understandings of the rubrics used are all factors which affect the reliability of these different/alternative assessments. The tasks themselves must be analyzed for the content and skills that are necessary to fulfill the task description. Designing tasks with specific understandings/competencies in mind and carefully analyzing the tasks for how they demonstrate particular content and skills is the foundation for developing reliable different/alternative assessments that can be compared over time as assessing the same domain. If the products that are evaluated are complex, created by students over time, multi-dimensional and personal expressions of the task, it is likely that the products will be more reliable indicators of what students understand than any one-shot sit-down test. Moreover, basic principles of different/alternate assessment methods require evaluating multiple examples of student work, assessing work longitudinally and focusing on growth and learning in the evaluation process. The composite results of these different/alternative assessment procedures yield a

---

[22] Gertrude Conlan, "'Objective' measures in writing ability" in K. L. Greenberg, H. S. Wiener and R. A. Donovan (Eds.) Writing assessment: Issues and strategies (White Plains, NY: Longman, 1986), p. 110.

[23] We noted earlier the numerous factors that influence test results. In a recent test in Massachusetts, seventy percent of students in one large school district did not respond to any of the open-ended items, which comprised approximately 30% of the overall score. Overall scores were clearly affected. This is a motivation issue.

very detailed picture of what students understand and know how to do. They also get close to students' attitudes, dispositions and motivation and, therefore, have the potential for being particularly reliable as a measure of student accomplishment. No score on a standardized test can provide anything close to this.

Validity, possibly the most critical aspect of tests, refers to the degree to which the test actually measures what it says it does, how close the connection is between the test and the knowledge or skills being assessed. If, for example, we were to ask students to correct spelling and punctuation errors in isolated sentences, we might infer something about their knowledge of the spelling code and sentence mechanics. But it would not be wise to infer a student's writing ability from such a skill audit. The complex act of communicating ideas in writing is not mirrored in this sampling of a basic skill. Validity, therefore, concerns the inferences we can credibly make about the test-taker's knowledge or ability based on his or her performance.[24] Establishing validity is difficult because it cannot be determined by statistical methods alone. Rather, content validity relies on the evaluation of the match between the assessment task and the ability or concept it is meant to tap. In the case of standardized tests, validity of test items to assess the appropriate knowledge and skills for students is, as we noted earlier, generally established by "experts" and agreed to by convention. But this is highly contested territory. Who should determine what the content and purpose of an education should be? Test-makers? Policy-makers? Legislators? Business leaders? Parents and community members? Teachers? Students? The Rural Trust believes that such matters are best decided locally, that validity is made more understandable at this level.

Another way that standardized tests establish validity is to compare test results across a number of instruments and measures. This is called concurrent validity. Once the content or construct has been determined, the degree to which tests purporting to assess the same content or construct produce similar results determines concurrent validity. This form of validity is also technical, not rooted in the actual relationship to what students can do or understand.

Different/alternative assessments face similar challenges and must adopt practices that permit valid inferences. Advocates of different/alternative assessments suggest several procedures to address validity concerns, such as constructing assessment tasks as analogues of real-life problems, analyzing tasks according to the competencies tapped and knowledge requirements, and creating scoring guides or rubrics that differentiate levels of performance systematically and appropriately. In addition, they understand that multiple data sources from a variety of tasks are necessary to adequately infer abilities and understandings, that no single task or test can tap the range of abilities and know-how we value as ends of education. The fact that projects used for assessment purposes are completed over time, are filled with complexities, and have embedded in them multiple products or tasks provides greater confidence about their validity.

In assessing the validity of tests and alternatives to tests, we need to frame our questions in terms of what the intended goal of the assessment or test is. If tests are used to audit learning, differentially reward students, label students according to skills and/or deficiencies, consider future life opportunities, the strictest scrutiny of content validity and reliability of results has to be demanded.[25] If we use assessments primarily to inform and adapt teaching and learning, we should be no less rigorous in constructing adequate tasks and evaluation practices.

Designing Different/Alternative Assessments

In discussing the design of different/alternative assessments, we think first of classroom use, informing students and the ongoing teaching-learning exchange. The better the assessments, the more useful they are -- not only for classroom use but for use in school-wide assessment as well.

---

[24] Many teachers have shared with us that their most accomplished readers often score poorly on standardized reading tests and their multiple choice technology. Deborah Meier, in Reading failure and the tests (New York: Workshop Center, City College of New York, 1974), found reading scores so unrelated to the actual reading abilities of her students that she questions whether the reading tests test reading. She is essentially raising a validity concern.

[25] If we actually applied such scrutiny, we would likely back away from all such uses.

Not all the different/alternative assessments we have seen meet the criteria necessary to function as educative tools, information generators. Many practices that go by the name of different/alternative assessments are not carefully planned and executed. They can be laborious exercises that offer little new information to teachers or students. Moreover, there are many classrooms where portfolios are merely collections of all of a student's work, unanalyzed for what the work tells about a student's growth. If the work in the portfolio is the collection of worksheets that require simple fill-in answers or uniform responses to unchallenging assignments, for example, it will be no more useful as an instructional tool than a grade book or a test. In order to create different/alternative ways of assessing what we value and desire students to learn, we need to be mindful of designing assessments that match well our purposes. Stated simply, good assessment practices require that teachers :

- Clearly articulate goals of instruction (What should students know, understand and be able to do? The tendency is to outline long lists that guarantee more coverage than uncoverage. Thinking about smaller numbers of generative goals is more productive.)[26]
- Determine valid and appropriate evidence for meeting instructional goals (What evidence will demonstrate student understanding and know how?)
- Establish criteria or standards for assessment (What level of competence is expected? What tasks inform us this level has been met?)
- Design assessment task/s with the following considerations:
  - Multi-dimensional, demonstrating complexity of understanding, know how (What does the task require in terms of knowledge, thinking, social competencies, personal habits of mind, etc.?)
  - Mirrors real-life task or problem to solve (What genres, contexts, and content are typical of real-world work in this area?)
  - Allows for multiple data sources (What pieces of evidence will be evaluated pertaining to the completion of the task?)
- Construct rubric/s for assessing product/s (What counts as excellent performance on the various components required by the task? What differentiates levels of performance from novice to expert?)
- Engage students in the evaluation criteria for quality performances
  - Involve students in developing the criteria articulated in the rubrics that will be used for evaluation
  - Observe the carrying out of the task and provide descriptive feedback according to mutually agreed upon criteria (What does the student/s work demonstrate in the terms used in the assessment rubric?)
  - Provide students with exemplars to demonstrate various levels of meeting the criteria
- Prepare evaluators in order to achieve reliable results
  - Have multiple evaluators evaluate completed products or public exhibitions, providing ratings for all the rubric categories
  - Discuss ratings among evaluators while they are engaged in the process (What do they use as criteria to justify a given rating? What discrepancies are there among different evaluators' initial evaluation scores and on what basis can they reach consensus?)
  - Insist on a high rate of agreement among evaluators, clarifying rubrics if necessary

---

[26] It might be helpful to consider the statement of purpose guiding the Central Park East Secondary School:

At CPESS we make an important promise to every student - - one we know we can keep. We promise our students that when they graduate from CPESS, they will have learned to use their minds - - and to use their minds well.

In every class, in every subject, students will learn to ask and to answer these questions.

1. From whose viewpoint are we seeing or reading or hearing? From what angle or perspective?
2. How do we know what we know? What's the evidence and how reliable is it?
3. How are things, events or people connected to each other? What is the cause and what is the effect? How do they "fit" together?
4. So what? Why does it matter? What does it all mean? Who cares?

Examples of Practices in Designing Different/Alternative Assessment Tasks

We now look more carefully at each of the recommended practices for designing good different/alternative assessments that can serve multiple purposes -- at the classroom as well as the school level. In addition to a description of each practice and its importance to the overall process, we provide several examples developed by a number of schools, including Rural Trust schools, as well as case studies of assessment practices. In undertaking the work of designing different/alternative assessments, it is useful for educators to learn from the many resources and examples available to them, even though they will ultimately have to develop their own. To begin with, we present a case of a Rural Trust site that is in the process of designing complex projects that are assessed across multiple dimensions. The three educators involved in this work have just finished their second year of implementing the projects described and are collectively considering next steps to enhance the work in the future. They are also struggling with better ways to assess their students' multi-disciplinary work.[27] We present this case study of the multi-media U.S. History productions at the Molly Ockett Middle School in Fryeburg, Maine to illustrate some of the challenges that many teachers and schools face as they work to effectively and formatively assess Rural Trust work. The discussions of purposes, planning and evaluation in which these educators are engaged also provide an example of the thinking that goes into making project work more accountable and more meaningful to the students.

Molly Ockett Middle School, Fryeburg, ME

On a hot, humid June evening, uncharacteristic of Maine weather in late spring, students, parents, teachers, community members and invited guests sat in the Molly Ockett Middle School cafeteria to view the video productions from Tom McLaughlin's eighth grade social studies class. The evening's program featured ten short computer/video documentaries that presented visuals, charts, graphs, computer art work and student-created narratives related to curriculum topics such as the institution of plantation slavery, immigration history in the U.S., and the Industrial Revolution. Each three to eight minute video production was created by groups of students who had formed into Production Companies for the purpose of completing a documentary or, alternatively, producing a newspaper of that time period. Even though the productions showed the novice attempts of eighth graders in multi-media transitions, coherence, and sometimes in regard to accurate historical details, the audience was uniformly impressed by these students' ability to command sophisticated multi-media software to produce sound and image documentaries as well as the research and writing skills that formed the backbone of the productions. Hearing the story of the behind-the-scenes work that went into the video productions made the event even more impressive.

Tom McLaughlin, eighth grade social studies teacher, Pat Higgins, the school librarian and technology maven, and Bob Cobb, former intern in Tom's class and currently a high school social studies teacher at the Fryeburg Academy, the private high school that serves as Fryeburg's alternative to public secondary education, worked together to introduce eighth graders to computer technology -- software that provided access to thousands of primary documents, photographs, film clips and maps from which students pieced together not only their productions but also their emerging understanding of several periods of U.S. history. With the help of these mentors as well as local historical institutions, students also researched how their town and region experienced these historical periods. Their productions included family pictures, oral histories of immigration, documentation of the impact of industrialization and railroads on life in Maine, along with some family stories of grandparents living through slavery or finding freedom through the underground railroad in the area.

A discussion with Tom, Pat and Bob revealed how the productions came about conceptually as well as physically. They spoke of their highest hopes for the project, their challenges, what they had learned from the year that will help improve next year's productions, and issues of assessment. How, they asked, can they reliably assess the complex work conducted by groups of students over a long period of time during class as well as in long after-school sessions in the library and in the classroom? How good does the production have to be to merit an A? How can individuals' contributions be assessed in such multi-layered projects? Which of the many aspects of the project should be assessed? Should all aspects be equally or differentially weighted? And, importantly, how well do

---

[27] Most teachers beginning such rigorous curriculum and assessment struggle for many years, refining their efforts over time. It cannot be otherwise as there isn't a prescription for this work. It demands considerable invention and re-invention.

students understand the history they are documenting? Do the productions capture and present those understandings?

Pat, Bob and Tom each expressed a different goal for the project. Pat was concerned that students learn to use state of the art technology well, be comfortable with a number of sophisticated applications and perfect their techniques of presenting different kinds of information simultaneously. She told us that "the kids take to technology so well. Last year, very few of them even knew how to scan. They learned that and much more in a year." She said they were preoccupied with sound -- music and voice tracks -- and particularly enjoyed the technology side of the work. Pat noted that, "their parents', grandparents' and great grandparents' jobs are shutting down -- the shoe factory, woolen mill and other local businesses that have been here for a long time are closing down. They need to have other options for work, for jobs." In addition to gaining understandings of technology, Pat agreed with Bob that some of the most important learning was interpersonal, learning to work together to accomplish a common goal.

After talking to local business operators about the proficiencies they felt students should demonstrate upon completion of school, Bob learned that they wanted to know if students "can take a project, time manage it, and work together to do it." Being able to work collaboratively and efficiently was a major concern. The idea of forming production companies, with the project manager generating an idea and "hiring" collaborators to work on the project, was born. Each production company was responsible for producing a multi-media presentation on some aspect of history. The curriculum of the class was divided into five major projects, each requiring research into the subject, creation of a product as well as a plan defining responsibilities and timelines for each person in the company. The project manager was responsible for making sure that the plan was carried out on schedule and if someone in the company was consistently missing deadlines or not fulfilling their responsibility, they were fired. The students who were "out of work" needed to "sell" their services to another company, or come up with their own project and hire others to form a new company and project. Not many students experienced being fired, and those who did quickly learned to assume more responsibility and make good on their commitments. Each new project required reorganizing into new production companies. By the end of the year, students had experienced five different production companies, mastered a number of specific technical skills, and learned to effectively work through teamwork problems that emerged.

Tom, the teacher of the class that required this work, was most concerned about the content learning that was taking place. He felt uninspired by the non-controversial, boring presentation of history that is common in textbooks. He had students compare the portrayal of historical events in their textbook with the documentation they uncovered utilizing primary source material. Tom introduced his students to the concept of "spin" and showed them how each seemingly neutral account of an event contained a bias that revealed the author's point of view. Students analyzed the work of Ken Burns for its subtle bias or point of view and examined newspaper accounts of the same event which revealed more blatant perspectives. Tom wants his students to understand their own biases and see that they are influenced by where they are born, their race, gender, age, and whom they listen to. He encourages them to recognize his biases and to become critical consumers of information about the past as well as what they read daily in newspapers or see on TV.

All of these goals need to be addressed in assessing the work. Assessment is an ongoing concern for Tom. He feels that this year of production company work was mostly successful -- they all learned their way through complex computer application problems with the help of on-line technical support and, sometimes, the software developers directly. Next year should run more smoothly since they now know how to "fix" many of the problems that confronted them this year. With more attention to the content of the productions possible as a result of fewer technical glitches, the issue of adequate assessment and, related to this, definitions of quality performances across numerous domains, take on greater prominence. Tom discussed how he had organized assessment practices this year and then, with his colleagues, brainstormed ideas that they would like to implement to improve student performances and possibilities for assessing them.

This year the assessment criteria were distributed to students in the form of Product Descriptors, lists of things that must be included in the productions and specifications of what the products should look like. Project managers were responsible for seeing that the product evolved in that fashion. The Product Descriptors for the video documentary and newspaper projects were as follows:

Product Descriptor

Documentary on Vital Links[28]

Parts	Points	Attributes
Length	20	Minimum 2 minutes - Shows clear beginning, middle, end - Reflects chosen position
Images	20	Minimum 10 from Vital Links - Reflects chosen position, smooth transitions
Sources	10	Included in credits
Music/Narration	10	At least two music pieces from Vital Links supporting position - music doesn't drown narrator; Clear, articulate, sufficient volume
Local History Angle	10	
Text	10	Easily deciphered, grammatically correct
Statistics	10	Smooth; image and sound synchronized
Whole production reflects pro or anti bias - or is balanced with two sides.	10	Easily detected by viewer

Product Descriptor

Newspaper on Vital Links

Parts	Points	Attributes
Masthead	3	In Old English font - size 60 at least
Headline	2	In a different font and reflecting point of view
Images	10	At least four from *Vital Links* (others may be from other sources) Reflect point of view
Articles	15	Factual, sourced, contemporary, 12 font
Columns/editorials	10	Reflect opinion, contemporary, 12 font
Statistics	10	Minimum 2 - Support articles/columns
Advertisements, cartoons, obituaries, etc.	10	Each counted in place of a column (must be approved in advance)
Local History	10	Approved by CEO, sourced, at least 1 image
Whole paper reflects bias - or is balanced	30	Point of view easily detected - 4 pages - correct grammar and spelling

In addition to these product descriptors, which guided the work, Pat Higgins utilized the following rubric to inform students of the elements that go into making a quality multi-media presentation.

[28] Vital Links computer software, published by Addison-Wesley, contains thousands of images, video clips, and historical documents that was used as the primary resource for the video productions.

Multimedia Project Rubric
by Dan Senstock, Winner Multimedia Mania, Middle Grades Division

Evaluator:_____ **Project Name:** _____
Students:_____

	Inc.	or	Avg.	Good	Excellent
Preliminary Work					
Students completed storyboards prior to computer creation.	----		5	7	10
Students used primary and secondary research which included factual and interesting information.	1		5	7	10
Students included a minimum of_____different resources.	1		5	7	10
Students created a running bibliography.	1		5	7	10
Students asked for permission of any copyrighted material.	1		5	7	10
Students focused topic into workable subtopics.	1		5	7	10
Design	1		5	7	10
Students showed evidence, presentation, synthesis.	1		5	7	10
Stack contains all elements ... text, graphics, sound, video, animation.	1		5	7	10
Stack contains all presentation elements ... introduction body, conclusion.	1		5	7	10
Students show design elements ... contrast between text and background, graphics, video are not obtrusive and enhance the presentation.	1		5	7	10
Multimedia elements are visible, adequate font choices, color schemes, sizes and styles are appropriate.	1		5	7	10
Navigation is easy and understandable.	1		5	7	10
Content	1		5	7	10
Information is relevant and interesting.	1		5	7	10
Students have used creativity in the design.	1		5	7	10
Students have used correct punctuation, complete sentences, grammar and spelling.	1		5	7	10
Students showed complete understanding of presented material.	1		5	7	10
Students used knowledge of file size manipulation to control the size of the presentation.	1		5	7	10
Students have labeled necessary video, graphics, maps for further understanding.	1		5	7	10
Students have used a majority of original artwork, animations, video, graphics, sounds, etc.	1		5	7	10
Presentation	1		5	7	10
The presentation is fluent from beginning to end.	1		5	7	10
Students understand presentation media such as LCD panels, projectors, sound equipment, video equipment, etc.	1		5	7	10
Students prepared and checked equipment prior to presentation.	1		5	7	10

These teachers thought that incorporating a similarly detailed rubric for the three aspects of the project -- technical expertise, historical content, and interpersonal work relationships -- would produce higher quality products in next year's class. Tom mentioned that since much of the work was done outside of school hours and outside of his classroom, he was less able to incorporate a timely review of all of the products for accuracy of information and depth of presentation than he would have liked. Even his attempts to critique student products with the whole class took far more time than he had envisioned. However, having this year's products to discuss with next year's students will provide a better starting point for making quality criteria explicit for each product. In Tom's words, "it was important to see what was possible. Now we can go into next year with even clearer expectations of what we can expect students to accomplish." Besides making explicit the dimensions of the project work via rubrics or detailed task descriptions, the group also thought of other ways to generate criteria for quality with the students. Students could, for example, analyze Ken Burn's Civil War documentary for quality criteria and apply them to student productions of the previous year to encourage self-evaluation. Students could also be required to develop resumés reflecting their work in each Production Company and use it to reflect on their growth and learning. Further, a human resources "job" might be generated to work with groups that have difficulties

BEST COPY AVAILABLE

mediating conflict or with individuals who are not working up to expectations. This would, they understood, require more focused attention to attributes of positive collaboration, time management, and responsibility -- habits of mind that are valued by the community.

All three teachers agreed that the best way to depict the stages of student work would be through a digital portfolio. While they haven't formalized the process yet, they want to experiment with rewritable compact disks so that students would be able to alter an ongoing portfolio year after year, throughout their school career. They believe it is important for students to see their own growth over the course of four years and keeping this type of portfolio could encourage them to engage in such self-assessment. In the meantime, however, the production company organization coupled with learning to use sophisticated technology is very popular with the students, their parents and community members. Students are learning to critically analyze and produce historical documents as well as develop a host of other skills -- from writing persuasive letters to Bob to enlist his services after school to learning how to manage complex, multi-dimensional projects within diverse groups. They are also reaching beyond the school doors to learn the local history from community members and local historical societies. And as these comments from members of the audience attest, the students are engaged. Pat commented, "kids would beg to stay after school to work on these [videos]." A parent said that, "instead of saying 'nothin' when I asked what [my son] had learned in school today, he would go into great detail about what he was doing with this project." A member of the State Department of Education and supporter of Rural Trust projects in Maine told Tom that his students had "learned the most history I've seen in a long time. Putting together documentaries to teach others is deep learning."

One can use the Molly Ockett Middle School story to frame the examination that follows of recommended practices for designing effective different/alternative assessments.

1. Articulated Goals of Instruction

Good assessment requires those in schools to first articulate the important goals of instruction. Articulation of the goals comes in answers to questions like these: What do you most want your students to be able to understand and be able to do at the conclusion of a unit or course or the end of elementary school or secondary school? What do you expect students to be able to do that they were not able to do before? (We often ask: What do you most want to say about your students when they leave this unit of study or course? When they leave this school?)

There is no lack of articulated goals of instruction available to schools and communities. National discipline-based standards, state frameworks, district guidelines, even textbooks, provide numerous lists of goals for instruction.[29] Given this bounty of suggestions and mandates, it is important for those in local schools and communities to discuss and come to consensus on what constitutes an education of value and how that is to be demonstrated throughout the school as well as in the actual work undertaken in each classroom. Rural Trust communities are increasingly defining purposes in terms of local circumstances and productive citizenship. Providing answers to fundamental questions help guide this process. The

---

[29] There are numerous resources available to educators to provide examples of instructional and educational goals. National organizations have published what they consider valuable outcomes in various disciplines: The National Council of Teachers of Mathematics issued *Curriculum and Evaluation Standards for School Mathematics* in 1989 and set out a blueprint that other discipline-based organizations have followed. Science goals have been developed by The American Association for the Advancement of Science in its 1989 document, *Science for All Americans: Project 2061* and the National Research Council, National Academy of Sciences, in its 1993 document, *National Science Education Standards*. In addition, The National Council of Teachers of Social Studies, the National Council of Teachers of Science, and the National Council of Teachers of English are resources for discussions and proposals for discipline-based standards. State curriculum frameworks have been, or are being developed, in most states as a result of federal pressure to provide accountability data in order to receive funding. They also serve as detailed resources for schools and communities that are defining and making public their own locally tailored educational and instructional goals. In the end, however, we believe the determination of what constitutes the critical purposes must be developed locally, put into language teachers, students, parents and community people actually understand.

following example of guiding questions and possible answers comes from ASCD's *Practical Guide to Alternative Assessment*:[10]

a. What Important Cognitive Skills Do I Want My Students To Develop?

I want them to be able to:

- Communicate effectively in writing, or more specifically, to write persuasively, to write good descriptions, and to write stories.
- Communicate effectively orally.
- Analyze literature using plot, character, setting, and theme.
- Analyze issues using primary source and reference materials.
- Analyze current events from historical, political, geographic, and multicultural perspectives.
- Design and conduct studies to aid decision-making about current or everyday problems.
- Use the scientific method.
- Use different media to express what they know.

b. What Social And Affective Skills Do I Want My Students To Develop?

I want them to be able to:

- Work independently.
- Develop a spirit of teamwork and skill in group work.
- Appreciate their individual strengths.
- Be persistent in the face of challenges.
- Have pride in their work.
- Enjoy and value learning.
- Have confidence in their abilities.
- Have a healthy skepticism about current arguments and claims.
- Understand that we all have strengths and that each person is able to excel in some way.

c. What Metacognitive Skills Do I Want My Students To Develop?

I want them to be able to:

- Reflect on the writing process they use, evaluate its effectiveness, and derive their own plans for how it can be improved.
- Discuss and evaluate their problem-solving strategies.
- Formulate efficient plans for completing their independent projects and for monitoring their progress.
- Evaluate the effectiveness of their research strategies.

d. What Types Of Problems Do I Want Them To Be Able To Solve?

I want them to be able to:

- Know how to do research.
- Solve problems that require geometric proofs.
- Understand the types of problems that trigonometry will help them solve.
- Apply the scientific method.
- Predict consequences.
- Solve problems that have no right answer.

---

[10] Joan L. Herman, Pamela R. Aschbacher, and Lynn Winters, A practical guide to alternative assessment. (Alexandria, VA: Association for Supervision and Curriculum Development, 1992), 25-26. We present this example to assist Rural Challenge participants in thinking about this aspect of assessment, even though we know that many schools and communities have already worked out their own purposes and instructional processes.

- Make healthy choices.
- Create their own unique expressions.

e. What Concepts And Principles Do I Want My Students To Be Able To Apply?

I want them to be able to:

- Understand what a democracy is.
- Understand cause-and-effect relationships in history and in everyday life.
- Understand the meaning of various logical propositions.
- Criticize literary works based on plot, setting, motive, and so on.
- Understand and recognize the consequences of substance abuse.
- Apply basic principles of ecology and conservation in their everyday lives.

These guiding questions can be answered according to local considerations and priorities that determine an education of value. But after providing our local answers, we are then faced with the challenge of asking: "What will cause us to say that our students 'understand cause and effect relationships...?'" The instructional and assessment tasks that get to that kind of issue can then be developed. Such a process is often called "planning backwards." Another question teachers can pose after delineating their largest purposes for the students is: "What topics and tasks are most likely to get my students to the ends outlined?" This is a recognition that content is not equally generative, always filled with large possibilities.

Much thought and discussion has gone on in Rural Trust sites across the country around the definition of a pedagogy of place. While the specifics of what each place values differ, there are some common elements that underscore this work at its best and most likely would be incorporated in some way into the instructional goals of Rural Trust schools.[31] Among these are:

- Ensuring a genuine reciprocity in school-community relationships.
- Engaging students in active learning relating to their specific locale and context.
- Embedding content in real-life (authentic) projects related to the community, its history, culture, economics and sustainability.
- Promoting a positive and lasting identification with the history, culture, people and geography of the rural community and its surroundings.

2 Determine valid and appropriate evidence of meeting instructional goals

This issue asks teachers to specify how they will know if students have learned what has been defined in the goals. (What would cause you to say that the students have come to understand what you most want them to understand?) The culture of traditional assessment methods would have us accept the math percentile score on a standardized test as proof of a student's math learning but, as we explained earlier, this provides inadequate evidence of student understanding. Well constructed different/alternative assessment practices that build in appropriate standards for evidence can offer more valid information about student learning and understanding. Let's take an example of a different/alternative math assessment and the standards for evidence of having met the instructional goals.

- Mathematical analysis: Students would demonstrate their thinking about mathematics and their mathematical analysis abilities by going beyond simple recording or reproducing of facts, rules, and definitions or merely applying algorithms.

---

[31] See the February 1999 report, Living and Learning in Rural Schools and Communities, Harvard Graduate School of Education, for a lengthy discussion and examples of curriculum and student work that illustrate the principles of a pedagogy of place.

- Mathematical concepts: Students would demonstrate understanding of important (specified) mathematical concepts by representing the concepts in different contexts, or making connections of these concepts to other mathematical concepts, other disciplines, or real-world situations.[32]

- Written mathematical communication: Students would demonstrate elaboration of their understanding, explanations, or conclusions in the form of diagrams or symbolic representations and prose that present convincing arguments.

- Habits of study and work: Students would demonstrate effective study and work habits by utilizing their time well, completing assignments on time, paying attention during class and group activities, showing initiative in taking care of themselves and others, taking an active part in class and group activities.

- Habits of social interaction: Students would demonstrate social maturity by being considerate of others' opinions and belongings, cooperating with others to achieve common goals, listening and speaking with consideration for others, showing self-control, being dependable and responsible.[33]

Given the numerous lists of educational goals that pertain to a student's knowledge base, habits of mind and social development, it is important, as we have noted, for particular schools to establish their unique instructional priorities. It is abundantly clear that all of the numerous curriculum frameworks being established by states and professional associations in each discipline area are impossible to support in any one class, or one student's education.[34] Choices must be made based on the priorities that emerge from dialogue concerning the role of schooling in each community.[35] Specifying these purposes and criteria at the local level, along with an articulation of instructional goals that respect local values, knowledge and contexts is the practice that infuses local priorities into what may be considered externally imposed agendas for school curricula. Many Rural Trust sites have had extremely productive community meetings around large purposes for their schools. Such a practice offers local communities and schools the opportunity to exert their influence in designing education for their children and young people. Combining local priorities with common standards of critical thinking, oral and written expression, and understandings of various principles and essentials in the disciplines can become the blueprint for an articulated pedagogy of place while at the same time ensuring that students will learn to demonstrate the habits of mind, knowledge and processes that are generally required for school success and academic opportunities beyond required schooling.[36]

---

[32] The concepts would, of course, be specified. In terms of articulating overall goals of instruction, this could be a compilation of the topics and concepts of the course, such as rational numbers, probability, geometric reasoning, variable equations, etc.

[33] This example of instructional goals is an elaboration of an example of Standards for Student Work in Mathematics presented by Fred M. Newmann, Gudelia Lopez, and Anthony S. Bryk, "The Quality of Intellectual Work in Chicago Schools: A Baseline Report." (Chicago, IL: Chicago Annenberg Research Project, 1998), p. 12. For this guidebook, their example has been modified to break it down into phases of planning and constructing an alternative assessment task in mathematics and non-content goals have been added.

[34] Researchers at Mid-continent Research for Education and Learning (McREL) calculated that it would take at least 22 years to get through the content called for in the various lists of standards. They present their findings in Robert J. Marzano, John S. Kendall, with Barbara B. Gaddy, Essential knowledge: The debate over what American students should know. (Aurora, CO: Mid-continent Research for Education and Learning, 1999).

[35] The Rural Trust has developed a Policy Statement on standards that supports standards developed at the local level. Copies are available from the Rural Trust Policy Program, P.O. Box 68, Randolph, Vermont 05060.

[36] The reality is that our purposes in schools and classrooms will always be an amalgam of what we understand to be definitions of well educated persons along with values that are particular to our local circumstances, beliefs and hopes. The problem is keeping the list of purposes small enough to accommodate intensive focus on the part of

These decisions about purposes are subject to change as priorities shift and as educators, parents, community members and students participate more centrally in the definition of an education of value. In short, this fundamental practice of visiting and revisiting educational priorities and instructional goals keeps the question of what constitutes an education of value in the forefront. All instructional, curricular, organizational, and assessment practices can then become a natural complement to the central, articulated, evolving goals.

3. Establish criteria or standards for assessment

Educators need to specify what quality of performance will count as having demonstrated the desired understanding being evaluated.[17] As in the case of instructional goals, there are many examples of standards and criteria that can be used for assessment. Generally, a statement of a standard assigns absolute criteria for determining a high quality performance in a particular disciplinary-related task or to more general instructional goals and habits of mind. These standards and criteria are typically aligned with the overall instructional goals of a school as well as with valid and appropriate evidence of what constitutes the meeting of those goals.

For example, standards defining the intellectual quality of student work in writing and mathematics developed for the Chicago Public Schools as part of the Chicago Annenberg Research Project, which we presented earlier, in part, states the following criteria:

Standards for Student Work
in Writing and Mathematics

A. Writing

Standard 1. Construction of Knowledge
Student performance demonstrates interpretation, analysis, synthesis, or evaluation in order to construct knowledge, rather than merely to reproduce information.

Standard 2. Disciplined Inquiry: Elaborated Written Communication
Student performance demonstrates an elaborated, coherent account that draws conclusions or makes generalizations or arguments and supports them with examples, summaries, illustrations, details, or reasons.

Standard 3. Disciplined Inquiry: Form and Conventions
Student performance demonstrates proficiency in grammar, usage, mechanics, and vocabulary appropriate to the grade level.

B. Mathematics

Standard 1. Construction of Knowledge: Mathematical Analysis
Student performance demonstrates thinking about mathematical content by using mathematical analysis; that is, going beyond simple recording or reproducing of facts, rules, and definitions or mechanically applying algorithms.

---

teachers and students. Most of the frameworks and standards developed in various states have become so large that they encourage a thin curriculum and considerable cynicism.

[17] We move here to the question of level or quality because concerns are moving beyond whether a student can, for example, write an essay to how well a student's essays are written. Our best hopes for our students must always be large.

Standard 2. Disciplined Inquiry: Mathematical Concepts
Student performance demonstrates understanding of important mathematical concepts central to the assignment; for example, by representing concepts in different contexts, or making connections to other mathematical concepts, other disciplines, or real-world situations.

Standard 3. Disciplined Inquiry: Written Mathematical Communication
The student's performance demonstrates elaboration of his or her understanding, explanations, or conclusions through extended writing; for example, through diagrams, symbolic representations, or prose that presents convincing arguments.

The authors note in their presentation of the above standards what they are intended to provide:

> Note that the standards for authenticity of assignments and student work do not attempt to prescribe the precise knowledge, skills, or dispositions that should be included in the curriculum. Reaching agreement on specifically what to teach in each subject and grade requires resolving professional and political disagreements that will occupy schools, districts, states, and professional organizations for years to come. But one virtue of the standards for authentic intellectual work is that they can be applied to the teaching of ... [writing and mathematics] content at any grade level. By defining the nature of intellectual work necessary for success in vocation, citizenship and personal affairs, the standards provide a common intellectual mission that can bridge otherwise divisive preferences of teaching different disciplines, different content within disciplines, or different groups of students (p. 12).

While we have presented the first three practices of good different/alternative assessment as separate steps, they necessarily inform each other, pushing educators to define specifically what is worth assessing and how they will identify what it is that they want to assess. The process must start with articulating instructional goals, but the clarity of the articulation will result in viewing these three steps together.

4. **Design assessment task/s with the following considerations:**
   - Multi-dimensional, demonstrating complexity of understanding and know how
   - Mirrors real-life tasks or problems to solve
   - Allows for multiple data sources

We now examine a sample of assessment tasks that meet the foregoing criteria and have been designed with specific instructional goals, concepts of appropriate evidence, and evaluation standards and criteria in mind. We begin with an extended but abridged algebra assessment task requiring multiple abilities and knowledge bases.

A PROJECT IN COST-EFFECTIVE DESIGN

Groups of 3-4 students become members of a design firm competing with other student firms for the right to design a large museum complex. The task requires that students solve a variety of problems which require algebraic concepts and skills related to rate, work, and rows/columns of seating. In addition, students must make some philosophical and aesthetic judgments about the content and form of the museum.

The complex will sit on a hill overlooking and due west of the city of Harrisville. The existing straight road leads back to the city, is 30 miles long, and has a continuous slope of 11 degrees on the three-mile stretch down the hill to the valley. Parallel to the existing road, all the way to the city, is a canal with locks and a railroad track. A new road will have to be built from the main road to the site. The vertical distance from the main road to the proposed site is 167

feet, according to survey maps.

The bidding calls for a museum/auditorium of 500,000 cubic feet. The building must be made of brick. (Assume that the roof and costs other than that of the walls remain constant in all possible designs; such factors as the size and cost of the windows and doors need not be considered for this exercise.)

The design must solve the following problems [this is a partial list]:

a. You have three types of brick to choose from, labeled A, B, and C. Each A brick is 2x2x1 and costs $1. Each B brick is 4x4x2 and costs $5. Each C brick is 4x2x1 and costs $3. The A brick factory is 5 miles away; the B brick factory is 3 miles away; and the C brick factory is 4 miles away. All three factories are accessible by rail, barge or road.

b. A truck can carry 1/3 the number of bricks carried by a barge up the canal. Both modes, used together, can move 3000 bricks a day in 2 round trips each. A train can carry all the bricks at once.

c. Rental costs for each mode of transportation (not counting labor): The truck rents for $50 dollars a day; the barge, $200 per day; the train, $3000 for the one trip it needs.

d. The average worker (AW) can lay 200 bricks per day at $10 per hour for an 8-hour day. The skilled worker (SW) can lay 300 bricks per day at $15 per hour.

e. To finance the job, you will need to borrow money. You have your choice between rates of 8% compounded daily or a 12% flat rate. Figure the cost of borrowing the money based on the fraction of a year the project will roughly last (figure to the week only).

f. The rate of the current in the canal is 4 mph. A strong tailwind of 15 mph is usually blowing when traveling to the site from the city. The total round-trip time by truck is 50 minutes. It takes 18 minutes by car, travelling at 40 mph to get from the site to the city.

g. Each design team must "hire" the teacher as a consultant (3 brief, all-class presentations are "free"). Cost: $1000 per 10-minute segment. Teams can join together to "hire" the teacher as a means of reducing costs. Teams must plan ahead, however, submitting a schedule of times they wish to lock in with the teacher.

Design Competition Rules:

a. Your team will submit a written plan (with each person's contribution clearly noted) and make an oral presentation to the class (all members presenting). Voting on "design points" (see below) will be done by the entire class, with teams unable to vote for their own design. As a listener you will be expected to accurately check the facts and figures presented.

b. The museum will house portraits and two artifacts each of the ten greatest mathematicians of all time. Explain your choices and justify their inclusion.

c. Include with all figures accurate graphs of the top and side views of the building and the hill-site with road.

d. Be prepared to justify figures in your presentation to the class. An inability by any one of your team members to explain a figure or cite the correct formula will cost one design point.

e. The three (or four) of you must finish a proposal by the last class of the week. Each day late costs you five design points.

f. Each mathematical error (either in presenting or checking others) costs you two design points. Omission of a necessary formula or problem solution costs you three design points.

g. Five design points can be earned by winning each of the various segments of the competition: aesthetics, completeness, effectiveness of presentation, most balanced use of team members, cost-efficiency in transporting costs, lowest wages, shortest construction time, lowest overall accurate bid.

This lengthy assessment meets all of the design criteria stated above. The students must work out a multi-dimensional problem applying their knowledge of algebra, mathematics as a discipline, real-world knowledge about buildings and sites, ability to work independently as well as part of a group, set priorities and organize their work. They must also demonstrate an ability to diagram, write clearly, completely and persuasively and effectively present their project to the class orally. The task simulates a real world task in cost effective design but provides the information that would need to be collected and researched in real life. Students are invested with the responsibility of presenting their groups' design as well as evaluating those of others. They also need to seek specific consulting from the teacher, making the students designers of their instruction and guidance within the parameters of the problem. And finally, the complex task requires multiple sources of data to be evaluated. The relative weight of any of the dimensions of the project can be determined in order to assign a grade, but the firmer criteria for quality performance in each of the domains of assessment are determined by established standards of performance and specified in a scoring guide or rubric (which will be discussed below).

This example relating to cost effective design shows the potential richness of authentic assessment tasks, but all tasks need not be that complex or lengthy. Another example of a math assessment complies with the same design criteria while requiring a simpler, though not one-dimensional product.

8th Grade Mathematics Performance Task:[38]

Assignment: Your group is going to design tiles which can be used to decorate part of the classroom. You can use shapes on the attached Shape Sheet to cut out as many copies of each of the shapes you need. You may use any other tools you wish (calculator, rulers, glue, string, protractors, compasses, pens etc.)

> [Accompanying this task was a list of vocabulary words that included *regular polygon* (a shape whose sides are all the same length and whose corners all have the same "sharpness"), *complex polygon* (a shape which is not regular), *tile* (a regular or complex polygon which is used like a puzzle piece to attempt to cover a surface), *tessellate* (to cover a surface with tiles, all the same shape, so that no tiles overlap and so that there are no gaps between tiles), *good tile* (a tile that tessellates), *bad tile* (a tile that doesn't tessellate).]

The task included two parts. For Part I, students were told to:

- Find out which regular polygons make good tiles (remember: good tiles tessellate). For each good tile you find, cut out enough shapes to cover half a page of paper to show that the tiles tessellate (you can also do this by tracing). For two shapes that are bad tiles, cover half a page showing how they overlap or leave gaps.

- Find a pattern that shows which regular polygons are good tiles. Write an explanation as to why these are good tiles. Based on the pattern that you have found, are there any other regular polygons which make good tiles? Why or why not? Write an explanation which uses information on the paper that you covered with the tiles and the patterns that you looked for.

Part II of the assignment asked the students to create complex polygons and explain how they could be used to cover a surface. Students were asked to write up their findings and conjectures about this. .

---

[38] Reproduced in Newmann, et al.. "The Quality of Intellectual Work in Chicago Schools," p. 25.

The foregoing assignment asks for a demanding mathematical analysis by requiring students to discover which geometric shapes will tessellate to cover a surface. The task asks them to reproduce their reasoning in a written commentary and simulates a real-life task.

The following complex assessment task requires students to apply their mathematical knowledge in demonstrating their conceptual understanding of change. The notion of change is applied across a number of disciplines. The task description also includes judging criteria that distinguish the quality of the performance. (The value of sharing this information with students, up front, will be discussed in detail below.) This is followed by a Social Studies/Writing example.

Individual Project: PREDICTING CHANGES[19]

The petals of a growing flower, the national debt, a fad sweeping the school, a contagious disease within the population, the temperature of a cup of coffee, the diameter of a tree; these and many other things change with time. Some increase in size or get larger in quantity. Others shrink in size or get smaller.

ASSIGNMENT

Select something that interests you that either grows or shrinks over time. Use the methods of this unit to make predictions about changes in the future. Before you begin actual work you need to check with your teacher and explain your ideas for the project to see if it is acceptable. Your project must satisfy the following criteria:

1. It deals with something that increases or decreases over time;
2. The change is quantifiable; it can be measured or counted;
3. There is a reason for making a prediction of future changes;
4. An audience exists for whom the predictions are relevant;
5. The mathematical methods studied in this unit can be used for making the predictions; and
6. The predictions are based on real data that you either look up or gather yourself.

Your report should be written for an audience such as a company, government agency, research institute, other math students, etc. In addition to using several methods to make predictions, your report should also include a discussion of how appropriate each prediction method is for the situation and some implications of what might happen if your predications come true.

JUDGING QUALITY: Mathematics Content

[*The following is one of three rubrics used to score this project. The other two are* Processes *and* Presentation of Report.]

Well Done

Add-on and multiply-by methods of prediction are used and compared. Graphs of each are shown. Report describes events that may make predictions invalid. Estimates of probabilities are included. Effects of each event on predicted numbers are considered. Mathematics is careful, accurate and appropriate. Many or all possible variations are considered. Multiple representations (diagrams, words, formulas, etc.) are used. Analysis is thorough.

Acceptable

Mathematics is generally careful and accurate, but with some errors in the work or in the logic applied. Limited number of possible variations are considered. Single representation is used. Analysis may need more care.

Revision Needed

Work may be the beginning of an acceptable report, but mathematics needs careful review and correction or refining.

Restart/Reteach

Quality of work needs improvement or the topic is not useful -- the project needs to be started over with further instruction.

---

[19] This assessment task and rubric is taken from Growth-Patterns, Equals, available from Lawrence Hall of Science, Berkeley, CA 94720.

Part of a Civil Rights Unit[40]

Attached are five newspaper articles that appeared regarding the Montgomery, Alabama Bus Boycott that followed the arrest of Rosa Parks.

• • • • • • • • • • •

The point of this paper is to write an informed opinion - an analysis – which evaluates press coverage of the protest.

As someone who has studied the Montgomery Bus boycott and read many first-person accounts of people who were participants, you have your own outlook about what occurred. Your individual view can, and should be, included; however, it shouldn't overwhelm the analysis you submit.

To prepare the paper you need to:

(a) Read each article carefully and highlight or underline important points.
(b) Summarize the situation that the articles address in one paragraph.
(c) Consider what each reporter did to write the story that appears. Who was interviewed, listened to, read about, observed by the reporters?
    1. Where would you guess each reporter was during the protest?
    2. What would you guess each reporter feels about:
        ...the protest itself?
        ...what caused the protest to be called?
(d) Select quotes which can be effectively used and which are not repetitive.
(e) Interview at least two individuals who were alive at the time of the Montgomery Bus Boycott about their remembrances of the protest as well as press coverage. (You might show them the newspaper articles you are working with.)
(f) Write a transcript of an interview you have had with one of the reporters. (Think of yourself as a participant in the protest engaged with one of the reporters of a newspaper account you have read.)
(g) Pay attention to the headlines used in each story. Consider the impact they have on the reader.
(h) Study photographs. How might the photographs which were used influence the views a newspaper reader gets of the event?

Keep in mind as you work on this paper that you are using multiple sources to develop an overall view of the press coverage. You might want to compare different papers or discuss sources that reporters used to write stories. Whatever you decide, you should constantly keep in mind that you need to support your statements effectively.

Your concluding analysis should not be an afterthought; it should be a key part of your paper. Begin the concluding section (not the concluding paragraph) of your paper with the words, "I think the press coverage of this event was..."

In this section, state your point of view in a way that selects out and emphasizes the key points from all the points you've been describing. This section should make your conclusions clear. You should cite key pieces of evidence to show why your conclusions are valid. (This will also be presented orally.)

When you think you have finished, proof-read your paper carefully. Then, have someone else read your paper and have that person sign it, saying that it reads well.

In preparing your essay you need to:

Highlight or underline the attached articles
Jot down questions and comments, ideas and points to include as you think and read
Organize your jottings into a structure that you plan to follow in your essay
Write down your interview questions (this can happen during or after your interview)
Consider your interview transcript

Please attach all these items to your final essay when it is submitted.
If you'd like help with this assignment, see me.
This assignment needs to be completed no later than the due date, as there will be a follow-up assignment.[41]

---

[40] Adapted from an assignment given to students at the Urban Academy in New York City.

BEST COPY AVAILABLE

This preceding social studies/writing assessment task has many aspects of a traditional, albeit complex, assignment. The product requires analyzing original documents which are then to be synthesized to construct the product. Both oral and written communication is required of the students.

The last example we present comes from a Rural Trust site, is complex, intended to focus attention on a local issue with a local audience. It calls for group as well as individual work, written as well as oral presentations.

9th Grade Science Energy Unit

Community Action Initiative Assignment

Noble High School, Berwick, Maine[42]

In this assignment you will identify an energy-related problem in Maine and develop a plan to help address this problem. There are two parts to the assignment.

PART ONE: CREATING AN ACTION PLAN

You will be expected to turn in a formal write-up of this part of the assignment. Each member of your group will need a copy for their portfolio.

Create a plan of action which includes:

- a clear statement of the problem you will address
- a description of what you see as a solution to the problem
- a plan that will help to achieve the desired change. This is the "action" piece. You should include what you want people to do to solve the problem and how you will encourage people to take the action you want them to. Remember that we would like to be able to implement some of these plans, so keep them as realistic as possible.
- A list of potential problems your group may encounter and need to overcome to achieve your goal. (Examples: attitudes, habits, values, costs, etc.)
- A list of key people or resources that could be helpful to you in getting your initiative off the ground.

Responses to these requirements need to be written neatly or typed and in complete sentences.

PART TWO: CREATING YOUR COMMUNITY ACTION INITIATVE

This portion of the assignment is your chance to begin to put your plan into action by creating an advertisement that will convince others to adopt your plan.

Your advertisement can take the form of a poster, brochure, full-page newspaper ad, or radio advertisement. Your ad should achieve the following:

- Promote understanding of the nature of the problem
- Convince the community of the severity of the problem
- Promote understanding of potential solutions to the problem and of what you want people to do to solve the problem

---

[41] At the Urban Academy, many tasks such as this one are followed up with yet a more complex task. Everything builds on what went before.

[42] This assignment was furnished by Liza Finkel.

5. Construct rubric/s for scoring product/s

A major element of evaluating complex authentic performances of understanding and know how is making sure that the assessments reflect a valid representation of students' work. Any knowledgeable evaluator (including students themselves) should be able to assign an appropriate rating or interpret the given rating in a stable, "objective" manner. In order to do this, there must be a clear description of what each rating or performance level entails against which the work has been evaluated. This description is most commonly referred to as a rubric -- a guide which depicts the criteria which have been decided upon for the assessment. A well constructed rubric for a particular assessment task requires specifying the following elements:

- What criteria will be used to evaluate a particular performance?
- What constitutes a range of quality in performing a given task?
- How we can distinguish between different levels of quality, the key criteria to focus on?

A well designed implementation of rubrics to evaluate student work also includes practices that relate to validity, reliability and fairness. Frequently, rubrics are accompanied by exemplars of performances that illustrate each of the levels. This helps orient evaluators to fixed criteria.[43] The exemplars are also helpful to students. In many cases, students don't do good work because they aren't sure what good work looks like in practice. We now take a more detailed look at each of the elements related to well constructed rubrics.

Rubrics specify performance criteria at differentiated levels of performance. Some rubrics use scales defining four or five differentiated levels corresponding to "expert/advanced," "proficient," "satisfactory/needs improvement," or "novice" performance of a task. While we prefer smaller numbers of performance levels, from three to six, some rubrics contain a wider range of performance differentiation, up to 10 levels. The number of levels specified, however, is not nearly as crucial as the explicitness of criteria used to differentiate performance levels, a fair representation of expected quality as well as consideration for development in achieving quality. In this way, rubrics are unlike other common grading or rating practices.

Teacher assigned grades on tests are often based on a number score which, as in the case of standardized exams, is set at an arbitrary number of correct or incorrect responses. What any particular grade means, therefore, is unique to the teacher and/or the assignment and is influenced by comparing work across students in a given class or grade. Rubrics, on the other hand, define what a quality performance looks like no matter who performed it. Once the highest standard of quality has been determined, less expert performance levels are defined to fit the number of levels decided upon by those designing the rubric. An effective rubric will take into consideration the developmental level of the performer. Quality writing in the first grade will be substantially different than in the 10th grade even though both rubrics may include phrases such as "clearly presents ideas in an organized fashion." Including examples of quality performances with the rubrics help anchor the criteria at an appropriate developmental level.

Along with the number of performance levels that a rubric might distinguish among, there is also a choice in how many criteria are to be included and how they will be evaluated -- individually or holistically. Three basic types of rubrics, from most to least particular, are particular trait rubrics, holistic trait rubrics and analytic rubrics. Particular trait rubrics single out one or two elements in a complex

---

[43] In an effort to develop reliable rating systems for the assessment of student progress as learners, fixed criteria for scoring have taken hold. While we present them here in great detail, in part because they are the standard technology used with different/alternative assessment, we want to offer a caution -- that they not absolutely limit what is seen in the student work. Given the different ways that students express themselves, it is also important to see beyond the rubrics as a way of distinguishing the uniquenesses that exist, to be in a position to support students and their learning most fully.

performance to evaluate and basically ignore other elements or criteria of the performance for assessment purposes. Analytic trait rubrics attempt to account for all of the criteria or standards that are demonstrated in a performance and allow for individual ratings of each of the basic traits. Holistic rubrics, on the other hand, merge all of the criteria into holistic descriptions of various levels of performance quality.

The best way to illustrate these various characteristics is to examine samples of rubrics that have been developed and notice what they choose to highlight in evaluating student work.

The following rubric was designed to evaluate a Senior Essay which required students to take a position concerning an article dealing with substance abuse. This particular rubric uses a 9 point scale. The criteria are specific to the task, narrowed to organization and style, making this a *particular trait rubric*. The person involved in the assessment is to assign a single score based on these specific elements. The criteria necessary for determining the level are carefully defined and one example of an indicator is provided to illustrate how a student would meet or not meet a particular criterion.

Rubric for Heritage High School, Littleton, Colorado, Senior Essay

9-8 The upper-range responses satisfy the following criteria:

a. *Summary.* The summary should identify the main idea [of the reading].
b. *Focus of agreement.* Agreement and/or disagreement may be complete or partial but writer must make clear what he/she is agreeing/disagreeing with. Specifically, 9-8 papers must address author's thesis, not substance abuse in general.
c. *Support for agreement/disagreement.* Support should provide an analysis of the argument and/or relevant and concrete examples.
d. *Style and coherence.* These papers demonstrate clear style, overall organization, and consecutiveness of thought. They contain few repeated errors in usage, grammar, or mechanics.

[The four phrases in italics represent the dimensions being scored. Two of the criteria are underlined.]

7 This grade is used for papers that fulfill basic requirements for the 9-8 grade but have less development, support, or analysis.

6-5 Middle range papers omit or are deficient in one of these four criteria:

a. *Summary.* Summary is absent or incomplete, listing only author's thesis.
b. *Focus of agreement/disagreement.* What the writer is agreeing/disagreeing with is not clear or is unrelated to the author's proposals. Example: writer doesn't use enough phrasing like "on the one hand'... on the other hand..." [an indicator].
c. *Support.* Writer only counterasserts; examples are highly generalized or not distinguishable from examples in the article. Analysis may be specious, irrelevant, or thin.
d. *Style and coherence.* These papers are loosely organized or contain noticeable errors in usage, grammar, or mechanics.

4 This grade is used for papers that are slightly weaker than the 6-5 papers. Also, a student who writes his/her own parallel essay in a competent style should receive a 4.

3-2 These papers are deficient in two or more of the criteria. Typically, they weakly paraphrase the article or they have serious organization/coherence problems. Papers with serious, repeated errors in usage, grammar, or mechanics must be placed in this range. [This whole paragraph, like all the previous ones, is a descriptor for this point on the scale.] [44]

[44] Reproduced from the example provided in Wiggins, *Educative Assessment*, p. 155.

The following is a general, *holistic rubric* for writing developed at Bronxville (New York) High School.[45]

Writing Assessment Criteria

[4] Exemplary/Excellent
- Good use of accurate details (intelligently selective)
- Interesting -- appeals to a wide audience
- Clearly makes its points, doesn't run-on or ramble
- Appropriate vocabulary / Good diction
- Correct mechanics, clear knowledge of mechanics
- Well-organized
- Smooth transitions

[3] Good
- Appropriate use of details (some lacking)
- Attempts to appeal to a wide audience; style is basically "smooth"
- Clear thesis, makes its points with an occasional run-on
- Vocabulary is adequate; little or no mis-use of words
- Occasional mechanical errors
- Generally well-organized; some problem with transitions

[2] Acceptable
- Misuse or lack of detail in significant places
- Transitions are weak -- or missing in places
- Narrow point of view, appeals to a limited audience
- Point may be distinct/clear BUT needs more support or details
- Organized but lacks focus, details (Doesn't clearly follow thesis at times)
- Limited vocabulary
- Mechanical errors

[1] Unacceptable
- Failure to use details or facts effectively
- Rambles, runs-on
- Poorly organized
- Simplistic vocabulary, misused or inappropriate diction
- Numerous mechanical errors
- No clear thesis or point of view

The value of holistic rubrics is their assistance in assessing overall performance on a task. As raters become familiar with the criteria and with the consensual assignment of scores to pieces of work, holistic rubric scores can be assigned quickly and students' work can be compared over time and for evidence of progress on general writing criteria. If, however, more detailed, or specific information about students' writing performances is desired, *analytic rubrics* can provide that fuller information. These rubrics identify separate traits that are combined in any written performance. Analytic rubrics require more than one rating on each piece of student work, but they also provide a more detailed indication of the quality of any student's writing. The more traits identified, the more details pertaining to writing ability can be tracked over time to assess student growth. Here are two examples of analytic rubrics for writing. The first rubric identifies six traits that go into effective quality writing. It was constructed around the following general descriptions, provided by teachers, that identified aspects of quality performance on each of the six traits:

---

[45] Johnson, The performance assessment handbook, p. 100.

SIX TRAITS FOR TEACHING/ASSESSING WRITING[46]

IDEAS
Clear -- makes sense
Writer has narrowed idea or story line to manageable proportions
Writer has plenty of information
A fresh, original perspective
Details that capture a reader's interest or make ideas understandable

ORGANIZATION
An inviting lead that pulls me in
Starts somewhere and goes somewhere
Even when the writer seems to drift from the topic, everything's connected
I have the feeling it's building to something
Doesn't just stop
Doesn't end "Then I woke up and it was all a dream."
Doesn't end with a redundant, banal, or preachy summary: "Now you know the three reasons why we must all join in the war on drugs."
A strong sense of resolution or completion

VOICE
Sounds like a person wrote it, not a committee
Sounds like this particular writer
Brings topic to life
Makes me feel part of it
Makes me feel connected with the writer -- maybe even want to *meet* the writer
Makes me respond, care what happens
Writer seems involved, not bored
Brims with energy

WORD CHOICE
Memorable moments
Words and phrases I wish I'd thought of myself
Word pictures
Every thought is crystal clear
Strong verbs
Simple language used well
Words used precisely
Minimal redundancy
The writer is *speaking* to me -- not trying to impress me

SENTENCE FLUENCY
Easy to read aloud
Inviting, playful rhythms
Well-built sentences
Varied sentence length and structure
Cadence

CONVENTIONS
Looks clean, edited, polished
Most things done correctly
Easy to decipher and follow
Free of distracting errors
Designed to make reading easy and pleasant
Attention given to spelling, punctuation, grammar and usage, capitalization, and indentation

The list of traits was then developed into the following rubric that identified levels of performance quality for each of the traits assessed for expository or narrative writing:

---

[46] This list as well as the resulting rubric which follows is explained in Vicki Spandel and Richard J. Stiggins, Creating Writers: Linking Writing Assessment and Instruction, Second Edition, (White Plains, NY: Longman, 1997) 49-57. It can be used, in whole or in part, for writing in all the subject areas.

Six-Trait Analytical Scoring Guide

Traits	4	3	2	1
Ideas	Essay is clear, focused, purposeful and enhanced by significant detail that captures a reader's interest. • The paper creates a vivid impression, makes a clear point, or tells a whole story, without ever bogging the reader down in trivia. • Thoughts are clearly expressed and directly relevant to a key theme or story line. • The writer selectively and purposefully uses knowledge, experience, examples and/or anecdotes to make the topic both understandable and interesting. • Quality details consistently inform, surprise, or delight the reader--or just expand his or her thinking	The writer has made a solid beginning in defining a key issue, making a point, creating an impression or sketching out a story line. More focus and detail will breathe life into this writing. • It is easy to see where the writer is headed, even if some telling details are needed to complete the picture. • The reader can grasp the big picture but wants more specific elaboration. • General observations still outweigh specifics. • There may be too much information; it would help if the writer would be more selective. • As a whole, the piece hangs together and makes a clear general statement or tells a recountable story.	The writing is sketchy or loosely focused. The reader must make inferences in order to grasp the point or piece together the story. The writing reflects more than one of these problems: • The writer still needs to clarify the topic. • The reader often feels information is limited, unclear, or simply a loose collection of facts or details that, as yet, do not add up to a coherent whole. • It may be hard to identify the main theme or story line. • Everything seems as important as everything else.	The writing is sketchy or loosely focused. The reader must make inferences in order to grasp the point or piece together the story. The writing reflects more than one of these problems: • The writer hasn't clarified the topic. • The reader often feels information is unclear or simply a loose collection of facts or details that, as yet, do not add up to a coherent whole. • It may be hard to identify the main theme or story line. • Everything seems as important as everything else.
Organization	The order, presentation, or internal structure of the piece is compelling and moves the reader purposefully through the text. • The organization serves to showcase or enhance the central theme or story line. • Details seem to fit right where they are placed. • An inviting lead draws the reader in; a satisfying conclusion ties up loose ends and leaves the reader with something to think about. • Organization flows so smoothly the reader does not need to think about it. • The entire piece seems to have a strong sense of direction and balance. Main ideas or high points stand out clearly.	The organizational structure guides the reader through the text without undue confusion. • Sequencing seems reasonably appropriate, given the main theme or story line. • Placement of details seems workable though not always deft. • Predictable moments or developments outweigh surprises or discoveries. • The introduction and conclusion are recognizable and functional. • Transitions are usually present but sometimes reinforce obvious connections. • Structure is sometimes so dominant it is hard for the reader to focus on the ideas or voice. • The piece has a developing sense of balance; the writer is zeroing in on what is most important but does not yet build to that point effectively.	Ideas, details, or events seem loosely strung together. The reader struggles to discover a clear direction or purpose. The writing reflects more than one of these problems: • There is as yet no identifiable structure to move the reader from point to point. • No real lead sets up what follows. • No real conclusion wraps things up. • Missing or unclear transitions force the reader to make giant leaps. • Sequencing feels more random than purposeful, often leaving the reader adrift in details or generalities.	Ideas, details, or events seem loosely strung together. The reader struggles to discover a clear direction or purpose. The writing reflects more than one of these problems: • There is as yet no identifiable structure to move the reader from point to point. • No lead sets up what follows. • No conclusion wraps things up. • Missing or unclear transitions force the reader to make giant leaps. • Sequencing feels more random than purposeful, often leaving the reader adrift in details or generalities.
Voice	The writer's passion for the subject drives the writing, making the text lively, expressive, and engaging. • The tone and flavor of the piece fit the topic, purpose, and audience. • Clearly, the writing belongs to this writer and no other. • The writer's sense of connection to the reader is evident. • Narrative text is open, honest, and revealing. • Expository or persuasive text is provocative, lively, and designed to prompt thinking and to hold a reader's attention.	The writer seems sincere and willing to communicate with the reader on a functional, if somewhat distant, level. • The writer has not quite found his or her voice. • The writer often seems reluctant to "let go" and thus holds individuality in check. • The writer has an idea of the audience but is reluctant to let them "in."	The writer seems somehow distanced from topic, audience, or both; as a result, the text may lack life, spirit, or energy. The writing reflects more than one of these problems: • The writer does not seem to reach out to the audience or to anticipate their interests and needs. • Though it may communicate on a functional level, the writing takes no risks and does not involve or move the reader. • The writer does not seem at home with the topic enough to personalize it for the reader.	The writing is purely formulaic. There is no evidence of personal involvement with the ideas or the story. The writing reflects most of these problems: • The writer shows no sense of an audience, only writing to fulfill the assignment. • The writer takes no risks and offers a bland treatment of the topic or story. • The writer seems to not care much about the topic.

BEST COPY AVAILABLE

Traits	4	3	2	1
Word Choice	Precise, vivid natural language paints a strong, clear, and complete picture in the reader's mind. • The writer's message is remarkably clear and easy to interpret. • Phrasing is original--even memorable--yet not overdone. • Lively verbs give the writing power. • Striking words or phrases stay with the reader, often prompting thoughts or insights.	The language communicates in a routine, workable manner; it gets the job done. • Most words are correct and adequate, even if not striking. • Energetic verbs or memorable phrases occasionally spark the writing. • Familiar words and phrases give the text an "ordinary, common" feeling. • Attempts at colorful language are there but not quite on the mark.	The writer struggles with a limited vocabulary, searching for words or phrases to convey the intended meaning. The writing reflects more than one of these problems: • Vague words and phrases convey only the most general sort of message. • Repetition clouds the ideas. • Words are used incorrectly. • The writer has a hard time grasping the intended message.	The writer shows little sense of appropriate or varied word choices. It reflects most of these problems: • Repetition distracts the reader. • Vague, general cliché's fill up the piece. • Words are used incorrectly. • The writer cannot readily grasp the meaning of the piece.
Sentence Fluency	An easy flow and rhythm combined with sentence sense and clarity make the text a good read aloud. • Sentences are well crafted, with a strong and varied structure that invites expressive oral reading. • Purposeful sentence beginnings often show how a sentence relates to and builds on the one before. • Sentences vary in both structure and length, making the reading pleasant and natural, never monotonous.	The text hums along with a steady beat. • Sentences are grammatical and fairly easy to read aloud, given a little rehearsal. • Some variation in length and structure enhances fluency. • Sentence beginnings help the reader's interpretation of the text.	A fair oral reading of the text takes practice. The writing reflects more than one of these problems: • It is sometimes difficult to tell when one sentence ends and another begins. • Ideas are hooked together by numerous connectives (and..but..so..then) to create one gangly, endless "sentence". • Short choppy sentences bump the reader through the text. • Repetitive sentence patterns grow distracting or bore the reader. • The reader must pause or reread to get the meaning. • Transitional phrases are missing or so overdone that they are distracting.	The text is very difficult to read aloud with any momentum. It reflects most of these problems. • It is sometimes difficult to tell when one sentence ends and another begins. • Ideas are hooked together by numerous connectives (and..but..so..then) to create one gangly, endless "sentence". • Short choppy sentences bump the reader through the text. • Repetitive sentence patterns grow distracting or bore the reader. • The reader must pause or reread to get the meaning. • Transitional phrases are missing or so overdone that they are distracting.
Conventions	The writer has excellent control over a wide range of standard writing conventions and uses them with accuracy and style to enhance meaning. • Errors are so minor that if there are any a reader can easily miss them unless searching for errors. • The text appears clean, edited, polished. • The text is sufficiently long and complex to demonstrate control of a range of conventions.	The writer shows reasonable control over the most widely used writing conventions and uses them with fair consistency to create text that is adequately readable. • There are enough errors to distract an attentive reader somewhat but they don't obscure the meaning. • It is easy to get through the writing without stumbling but it clearly needs polishing. • Moderate editing is necessary to get the piece ready to "publish."	The writer demonstrates limited control even over widely used writing conventions. Oral reading of the text takes practice. The writing reflects more than one of these problems: • Unclear writing, punctuation, or irregular word patterns make it hard to tell where one sentence ends and the next begins. • No visual evidence of paragraphs or erratic indentation. • Spelling is erratic; a number of common words are misspelled. • Extensive editing would be required to prepare the text for publication.	The writer demonstrates limited control even over widely used writing conventions. Oral reading of the text takes practice. The writing reflects most of these problems: • Unclear writing, punctuation, or irregular word patterns make it hard to tell where one sentence ends and the next begins. • No visual evidence of paragraphs or erratic indentation. • Spelling is erratic; a number of common words are misspelled. • Extensive editing would be required to prepare the text for publication.

The following analytic trait rubric is a simpler version to apply to writing because it combines criteria into only three large categories. One's preference for a more detailed rubric will depend on what the purpose of the assessment is, how many pieces of work need to be assessed and how evaluators determine the ease of applying any given rubric to examples of student work.

BEST COPY AVAILABLE

Three Trait Writing Rubric[47]

Content and Organization	Style and Voice	Mechanics
4 -Answers the questions or responds to the task completely and with details -Includes substantial information that is accurate and true -Elaborates using ample details, examples, reasons, illustrations, and other sources -States clearly what the purpose is -Stays focused on the topic -Has a clear beginning, middle and end, and is well organized, logical and sequential	4 -Uses a clear and consistent point of view -Tries something original, unexpected, uncommon or imaginative -Uses words in a way that creates pictures in the reader's mind (colorful words, and lots of action words) -Includes a variety of complete sentences -Is passionate, honest, personal -Keeps the reader interested -Uses rich and varied vocabulary -Uses tone that is appropriate for the purpose and audience	4 -Incorporates significant revisions -Is polished and presented in a legible and appropriate format -Follows rules of grammar, punctuation, spelling with very few errors -Uses correct subject-verb agreement - Uses correct verb tenses. -Uses a format that suits the assignment
3 -Answers the questions or responds to the task -Includes information that is accurate and true -Elaborates using some details, examples, and reasons -States the purpose -Stays focused on the topic -Has a clear beginning, middle and end, and is organized and logical	3 -Uses a fairly consistent point of view -Tries something original -Uses words in a way that creates pictures in the reader's mind (colorful words, lots of action words) -Is interesting -Uses some advanced vocabulary	3 -Incorporates revisions -Is polished and presented in an appropriate legible format -Follows rules of grammar, punctuation, and spelling with some errors -Includes some different and complete sentences -Uses a format that suits the assignment
2 -Partially answers the question or responds to the task -Includes minimal information -Uses minimal details, examples, and reasons with some errors -The purpose is unclear -Strays from the topic -Beginning, middle and end may be unclear -Is not always organized, logical or clear	2 -Point of view is unclear and/or inconsistent -Uses little descriptive language or colorful words Does not interest the reader Uses limited vocabulary	2 -Makes few revisions -End product is sloppy -Many errors in grammar, punctuation, and spelling -Product is illegible -Does not use the same format throughout the assignment -Uses few complete sentences
1 -Does not answer the questions or respond to the task -Includes minimal information that is sketchy or false -Uses few details or examples to support topic -Purpose is not stated -Strays from the topic -Shows no organization	1 -Shows no clear point of view -Uses little descriptive language -Is not interesting, incomplete -Uses very limited vocabulary	1 -Does not incorporate revisions -Has numerous errors in grammar, punctuation, and spelling -Is illegible -Uses many run-ons -Uses no format

[47] This writing rubric was developed at the Robert G. Shaw Middle School in Boston, Massachusetts.

BEST COPY AVAILABLE

The rubric below, while including many of the same elements we saw in the holistic and analytic rubrics above, serves as a less effective tool for reliably scoring student work. (We present it to enlarge our understanding of rubrics.) Since there is no explanation or definition of what constitutes and, more importantly, what differentiates the various evaluations (None, Entry Level, Workable, Competent and Excellent), the resulting evaluations for individual students will demonstrate the same weaknesses of letter grades.

	None	Entry	Workable	Competent	Excellent
MECHANICS OF WRITING Correct spelling.					
Grammatically correct Standard English sentences.					
Chosen vocabulary is appropriate to situation.					
Tone & style of writing are appropriate for the purpose.					
Ideas are developed into coherent paragraphs.					
Writing has been improved through restructuring, correcting errors and rewriting.					
STRUCTURE OF ESSAY Introduction clearly presents essay's ideas.					
Thesis statement is clearly identifiable.					
Each body paragraph performs an identifiable function.					
Topic sentences clearly relate paragraphs to thesis.					
Conclusion concisely summarizes essay's ideas.					

Since there is no way to fix the various levels on specific descriptive criteria, there is no guarantee that multiple evaluators are using the same criteria to differentiate among levels of expertise or quality. Therefore, there is no reliability in the assessment tool, making it impossible to draw valid inferences about the overall quality or ability of individuals or groups of writers. With this type of vague scoring guide, there will be little concrete evidence to show actual student growth or learning.

The array of rubrics to assess writing presented above demonstrates that there are numerous options available to teachers and schools. Also, what is selected will depend on the choices made concerning instructional goals, criteria and standards to be taught and met, as well as the function of any particular assessment.

Rubrics are being generated in all subject areas and at all grade levels for a variety of tasks and are reproduced as examples in numerous publications concerning new modes of assessment.[48] It is often helpful to start with examples of rubrics that have been developed by various teachers to fit the needs of their assessment practices. Others can learn from the language the rubric provides, compare the strengths and weaknesses of each rubric and then construct rubrics tailored to their own unique situations/goals. Each assessment should be conducted using a rubric specifically designed for that purpose. There are clear advantages to using very similar rubrics across assessment tasks in order to track student learning over time, but adding new criteria or modifying a basic rubric to fit the special needs of a particular

[48] The bibliography will be helpful in this regard.

assessment will not seriously interfere with the ability of the rubric to show student progress over time. Tailoring rubrics to specific assessment tasks provides better quality information to both teachers and students than force-fitting a single discipline-specific rubric on multiple assessment tasks.

We present below several rubrics that have been developed for a variety of tasks in a number of disciplines to help teachers consider ways of designing their own task-specific rubrics for assessments. The rubric which follows serves to score open-ended math problems.

Mathematics Problem Solution Generalized Holistic Rubric[49]

Demonstrated Competence

Exemplary Response ... Rating = 6

Gives a complete response with a clear, coherent, unambiguous, and elegant explanation; includes a clear and simplified diagram; communicates effectively to the identified audience; shows understanding of the open-ended problem's mathematical ideas and processes; identifies all the important elements of the problem; may include examples and counterexamples; presents strong supporting arguments.

Competent Response ... Rating = 5

Gives a fairly complete response with reasonably clear explanations; may include an appropriate diagram; communicates effectively to the identified audience; shows understanding of the problem's mathematical ideas and processes; identifies the most important elements of the problems; presents solid supporting arguments.

Satisfactory Response

Minor Flaws But Satisfactory ... Rating = 4

Completes the problem satisfactorily, but the explanation may be muddled; argumentation may be incomplete; diagram may be inappropriate or unclear; understands the underlying mathematical ideas; uses mathematical ideas effectively.

Serious Flaws But Nearly Satisfactory ... Rating = 3

Begins the problem appropriately but may fail to complete or may omit significant parts of the problem; may fail to show full understanding of mathematical ideas and processes; may make major computational errors; may misuse or fail to use mathematical terms; response may reflect an inappropriate strategy for solving the problem.

Inadequate Response

Begins, But Fails to Complete Problem ... Rating = 2

Explanation is not understandable; diagram may be unclear; shows no understanding of the problem situation; may make major computational errors.

Unable to Begin Effectively ... Rating = 1

Words do not reflect the problem; drawings misrepresent the problem situation; copies parts of the problem but without attempting a solution; fails to indicate which information is appropriate to the problem.

No Attempt ... Rating = 0

---

[49] This rubric was developed by the California State Department of Education, 1989, and is reproduced in Herman, et al. A practical guide to alternative assessment, p. 56.

BEST COPY AVAILABLE

The analytic rubric for science experiments below includes elements of oral and written communication as well as critical elements of the experiment.

Analytic Trait Rubric for Fifth Grade Science Experiments[51]

Experiment Design		*Scientific Results*	
4	Design shows student has analyzed the problem and has independently designed and conducted a thoughtful experiment.	4	Pamphlet explained with convincing clarity the solution to the problem. Information from other sources or other experiments was used in explaining.
3	Design shows student grasps the basic idea of the scientific process by conducting an experiment that controlled obvious variables.	3	Pamphlet showed that student understands the results and knows how to explain them.
2	Design shows student grasps basic idea of scientific process but needs some help in controlling obvious variables.	2	Pamphlet showed results of experiment. Conclusions reached were incomplete or were explained only after questioning.
1	Design shows student can conduct an experiment when given considerable help by the teacher.	1	Pamphlet showed results of the experiment. Conclusions drawn were lacking, incomplete, or confused.
Data Collection		*Verbal Expression*	
4	Data were collected and recorded in an orderly manner that accurately reflects the results of the experiment.	4	Speech presented a clearly defined point of view that can be supported by research. Audience interest was considered as were gestures, voice, and eye contact.
3	Data were recorded in a manner that probably represents the results of the experiment.	3	Speech was prepared with some adult help but uses experiment's result. Speech was logical and used gestures, voice, and eye contact to clarify meaning.
2	Data were recorded in a disorganized manner or only with teacher assistance.	2	Speech was given after active instruction from an adult. Some consideration was given to gestures, voice, and eye contact.
1	Data were recorded in an incomplete, haphazard manner or only after considerable teacher assistance.	1	Speech was given only after active instruction from an adult.

The next rubric, is specifically designed for the Heroes Exhibition Project which we presented earlier on pages 7-8. Two levels of performance are specified and the criteria speak specifically to the required elements of the assigned task.

[51] From Wiggins, Educative assessment, p. 167.

BEST COPY AVAILABLE

FULTON VALLEY PREP/PINER HIGH SCHOOL
HEROES EXHIBITION RUBRIC[52]
A Collection of Heroes

	Mastery	Distinguished
PROJECT CONTENT	1. The project contains a clear, thoughtful definition of a hero for our society. 2. The project includes a collection of heroes who fit the author's definition. 3. Each hero in the collection is described accurately and connections between the heroes and the definition are evident. 4. The author has researched the heroes in the collection and has included a bibliography of sources.	1. The project contains a clear, thoughtful, multi-faceted definition of a hero for our society. 2. The project includes a varied collection of heroes who fit the author's definition. 3. Each hero in the collection is described accurately and in detail and connections between the heroes and the definition are clearly drawn. 4. The author has thoroughly researched the heroes in the collection and has included a bibliography of sources.
PROJECT FORMAT	1. Written material is the original creation of the author or is clearly labeled as a quotation with the source identified. 2. Written materials are organized and readable, with an introduction, body and conclusion. Hypercard stacks allow easy travel among cards. 3. Written materials are neatly hand-written in ink or typed and contain spelling, punctuation or usage errors. 4. The layout and graphics in an anthology, calendar, cards, magazine or mural are neat and easy to follow. 5. The performer (series of monologues or speech) speaks clearly and loudly and, although using a script, shows evidence of ample rehearsal.	1. Written material is the original creation of the author or is clearly labeled as a quotation with the source identified. 2. Written materials are well-organized and interesting, with an introduction, body and conclusion. Hypercard stacks allow easy travel among cards and use graphics and backgrounds creatively. 3. Written materials are typed and contain minimal spelling, punctuation or usage errors. 4. The layout and graphics in an anthology, calendar, cards, magazine or mural are creative, eye-catching and professional-looking. 5. The performer (series of monologues or speech) speaks clearly and loudly and has memorized the material, referring minimally to note cards. The presentation is energetic, believable and convincing.
DEFENSE	1. The presenter can explain reasons for his/her definition and can field questions about the heroes in the collection.	1. The presenter can explain connections between his/her definition and our society and can thoroughly defend her/his choice of heroes for the collection.

[52] Presented in Johnson, The performance assessment handbook, Vol. 2, 132-133.

BEST COPY AVAILABLE

The rubric that follows can be used in relation to projects in any of the various subject matters. We have seen it used in history/social studies, science, drama and literature/language arts.

TASK: RESEARCH PAPER[53]

	OUTSTANDING	GOOD	COMPETENT
TOPIC	Presentation of a significant and complex issue	Presentation of an interesting issue	Presentation of an assigned question
SOURCES	Detailed evidence drawn mostly from primary sources: uses a wide variety of sources Evaluation of substantial number of conflicting secondary sources Brings analysis to a deeper level by employing additional, independent research in order to resolve analytic conflicts in the paper, i.e., disagreements between different sources	Some evidence drawn from primary sources: uses a variety of sources Evaluation of some conflicting secondary sources Works thoroughly with all available sources: carefully explores conflicts within the framework of the sources that are used	Adequate use of primary sources Adequate use of conflicting secondary sources Refers to and analyzes several available sources
ARGUMENT (organization)	An approved paper must adhere to all of the following argument-related criteria: 1. Acknowledge, and indicate understanding of, the readings and other sources. 2. Address the main issues raised by the readings. 3. Make clear and well supported arguments based on student's own point of view. 4. Explain and respond to conflicting arguments.		
	Complex argument clearly presented and supported by specific and relevant evidence: explanation of why opposing arguments are less valid	Presentation of clear argument supported by reasonable evidence: some opposing arguments countered, but not all	Argument evident, though organization may be somewhat weak: opposing arguments addressed in a more general manner
STUDENT VOICE	Confident writing style: student voice is evident: writes with lively, engaging language	Writing is clear and focused: style is straight-forward but not original	Writing is generally clear
CONNECTIONS	Clear, compelling introduction: conclusion that effectively synthesizes the strands of its main argument. All aspects of the paper support the overall structure. Clear, effective transitions. Establishes strong links to broader context.	Clear introduction and well constructed, thoughtful conclusion. Most aspects of the paper support the overall structure. Clear transitions. Establishes links to broader contexts.	Clear beginning and relatively persuasive conclusion. Most aspects of the paper support the overall structure. Makes appropriate though not always effective transitions.
CONVENTIONS	Grammar and punctuation flawless: appropriate documentation of sources: knows how to use quotations and paraphrasing to sustain an argument	Some grammar and punctuation errors, but writing is solid overall: appropriate documentation of sources: knows how to use quotations and paraphrasing to sustain an argument	Some grammar and punctuation errors: sources are correctly documented though occasional errors in paraphrasing and quotations.
EXTERNAL VALIDATION External reader/oral defense	Using the above Criteria	Using the above Criteria	Using the above Criteria

[53] This rubric was developed at the Urban Academy, New York City.

BEST COPY AVAILABLE

The final rubric we present in this set of examples evaluates aspects of student work that are not necessarily visible in written products, demonstrating that rubrics can be used to evaluate many aspects of student performances that are part of instructional goals. It is a rubric developed to assess a five-minute persuasive speech.

Rubric for Assessing a Speech[54]

Performance Task: Students will present a five-minute persuasive speech
Goal/Standard: Speak effectively using language appropriate to the situation and audience

Scale: **Criteria:**	**0** Not Yet	**1** Student Council Elections	**2** The Senate Floor	**3** Presidential Debates
Organization				
▪ Hook	None	Introduces topic	Grabs attention	Electrifies audience
▪ Transitions	None	Uses words to link ideas	Makes key connections between ideas	Smooth flow of ideas
▪ Closure	None	Lacks interest	Referred to introduction	Powerful and dramatic
Content				
▪ Accuracy	3 or more factual errors	2 factual errors	1 factual error	All information correct
▪ Documentation	No sources cited	1 source cited	2 sources cited	3 or more sources cited
▪ Quotations	No quotes	1 quote to support case	2 quotes to support case	3 key quotes to prove case
Delivery				
▪ Eye Contact	Reads speech	Looks at some people some of the time	Looks at some people all of the time	Looks at all of the people all of the time
▪ Volume	Could not be heard	Could be heard by people in the front	Could be heard by most people	Could be heard clearly by all people
▪ Gestures	None	Used a few gestures	Used some gestures appropriately	Used many appropriate gestures effectively
Visual Aid				
▪ Graphics	None	Minimal	Colorful	Creative graphics that enhance speech
▪ Appeal	None	Little visual appeal	Captures our attention	Visually stimulates audience
▪ Relevance	None	Minimal relationship to topic	Relates specifically to topic	Relates and reinforces topic

[54] This rubric appears in Kay Burke, Robin Fogarty, Susan Belgrad, The mindful school: How to assess authentic learning (Arlington Heights, IL: IRI/SkyLight Training and Publishing, Inc., 1994), p. 63.

Suggestions for Constructing Rubrics

Constructing rubrics for the first time can seem a daunting task. With so much to assess in well constructed, authentic, multi-dimensional instructional and assessment tasks, it might seem that the rubric needs to be similarly complex, but that doesn't have to be the case. Constructing rubrics begins with making choices. The choices to be made are dependent on the steps that preceded the task construction: determining the instructional goals and the criteria or standards to be assessed. These choices will determine the most salient dimensions of the task to be assessed and will necessarily limit the scope of the rubric.

The next set of choices pertains to the type of rubric that will best suit particular needs. Do you want to use a task specific rubric and identify the main items of the assignment specifically? Or do you want to use a more generic rubric that may apply to a number of assignments in that domain (for example, a generic math problem-solving rubric, an essay writing rubric, a science experiment rubric, an oral presentation rubric, an essential question rubric, and so on)?

When beginning this process it is helpful, as we have noted, to look at work that has already been done by other teachers across the country. Reviewing samples of rubrics, like the ones we have presented here, and adapting appropriate ones for particular purposes makes the initial process more manageable. Using constructed rubrics as models for the language, dimensions and differentiation of levels of performance can provide a good starting point for drafting your particular rubric. Tailoring it to your particular purposes, however, comes not only from your instructional choices; it also depends on the work of the students. Looking at student work together with other educators and discussing the performance of a number of students on a given assignment helps anchor the important dimensions of quality that are worthwhile assessing. These collaborative discussions of student work are important, not only for building consensus on what needs to be assessed, but also on what counts as quality work. They set the groundwork for creating reliable evaluations -- the same ratings assigned by multiple evaluators -- when a rubric is used for assessment.

The following scenario portrays a way to get started designing your own rubric:

A group of six teachers at Bloomfield High are meeting on a Saturday to design the rubric that will be used to assess the end-of-the-semester history exhibition -- a whole-class authored and illustrated booklet describing the post-WWII transition period in their town, using local history to illustrate some of the major societal transformations that occurred across the nation in the late 1940s and 1950s. Individual work on the project will be an oral history of a family member or local person who experienced the post-WWII changes in Riverview County. The students will then collectively produce a class-designed booklet that will tell the story of that historical period in Riverview County. The task description has asked them to:

- *provide the historical context for the individual stories*
- *organize the stories into logical units*
- *provide a brief commentary of the trends and transformations that the groups of stories illustrate*
- *mock up the layout of the booklet including the cover, title page, table of contents, illustrations and photographs, and the placement of text*
- *edit the volume and produce camera-ready copy which will be printed and distributed in the local historical museum*

The interdisciplinary teachers who are coaching the students on this project are meeting now to decide on the criteria that are the most important for evaluating this task. They listed all of the relevant criteria they could come up with on the board -- asking such questions as "what would characterize good work on this assignment?" They decided to concentrate primarily on the writing of history for the assessment. They will also evaluate the final product for aesthetic presentation and evaluate collaborative effort in producing this booklet at different times throughout the project. For now, though, they will work on the specific criteria relating to the writing of history.

John Hamilton, the English teacher, has brought in a number of rubrics relating to writing performances to use as resources. Susan Cawthorne, the history teacher, has provided the group with the list of criteria she shares with her students when they are assigned a research paper in history. In addition, she has brought samples of oral history projects from past years that demonstrate, in her opinion, high quality student work.

"Let's start this by seeing if we all agree with Susan that the work she brought is high quality," suggests Carlos Perez, the science teacher. "Not that I want to criticize the way you grade, Susan, but how do we know if we

BEST COPY AVAILABLE

see the same things that you do in the papers?" "Sounds like the right place to start," Susan agrees. The group quickly decides to holistically score the work on a 4-point scale: 4 is terrific, outstanding; 3 is very high quality, no real problems; 2 is still in the works, several things could be done much better; and 1 is back to the drawing board. There is no discussion as to what those labels mean for each person at this point; they just want to see how they each apply them to the same samples of work. Since Susan has brought what she considers to be only high quality work, they are expecting the scores to fall in the 3 to 4 range for the most part, although some "tough" graders will probably assign some 2s. The students' names have been removed from the papers, so the scorers are only looking at the work itself. They each put a sticky note on the back of the paper once they've read it with their name and score so that no one will see another's evaluation until they have multiply scored all the papers.

Once at least two people have read each of the 20 papers Susan brought, they look at the scores assigned and divide the papers into two piles -- those that received the same score from all evaluators, and those that did not. The same-score papers are then divided according to the assigned scores. The pile contains four 3s, one 4 and one 2. The pile of differently-scored papers is larger, but is put off to the side for the moment.

Susan goes to the board and agrees to write down the criteria that emerge from the group -- essentially what they based their scores on. They begin with the 4 paper, which they all agreed to be outstanding work. The list that results from the collective comments seems to fit into several categories or traits:

<u>Content</u>
Identifies the major social changes of the period
Frames the topic historically and personally
Topic is developed into several major points
Thorough treatment
Good illustrative details of major points
Logical inclusion of details
Original take on the subject
Informative
Picks effective details and quotes from oral history interview
Lets the reader get to know the person who is telling her story to the author
Great title

<u>Organization</u>
Introduction tells me where the essay is going
Major organization is chronological
General comment with specific example for each major idea
Paragraphs are topically organized
They flow from one to another smoothly with transition sentences
No gaps in logic or story
Conclusion wraps up the whole piece
Conclusion makes me relate the topic to today

<u>Style and Voice</u>
Engaging style
Rich, appropriate vocabulary
Varies sentence structure and length
Easy to read aloud, sounds good
Sounds like the author knows what s/he is talking about
Picks effective quotes from oral history interview

<u>Conventions and Mechanics</u>
Correct grammar, spelling, punctuation and formatting throughout

The group can readily see what each of these items means in terms of the essay they all scored as a 4. And they feel comfortable with these descriptions and categories. They now go through a similar process of description with the group of papers scored as 3s. There is little disagreement as to why the 3s are good, but not quite as good as the 4 paper and they can state the reasons in terms of the criteria they identified for the 4 paper. In fact, most of the comments are the same as those listed on the board, but there are a few areas that show less strength; in one case the author wrote in a less engaging style; the sentences weren't varied and the essay sounded simplistic. In

BEST COPY AVAILABLE

another, there were some quotes from the interview that didn't quite match the general point it was supposed to illustrate and there was no explanation about that. In yet another, the conclusion was basically a summary of the important points without any new reflection or connection to the present. The paper receiving a 2 was also clearly lacking in several critical areas: the introduction didn't provide important background information and there was only one good major point developed -- the others were given as examples with no discussion at the general level. Even Susan saw this as lacking in quality when she read it with the group, and admitted that she had been influenced by knowing so well the student who wrote it when she initially assigned it an A.

The next step in the process came after a break. They all realized that reconciling the pile of papers that were scored as both 4s and 3s or 3s and 2s, even one 4 and 2, by different evaluators would take some careful scrutiny to see what was at issue and how to reach consensus on differentiating between the two scores. The discussion they had about these papers after the break was rich and detailed, relying on specific examples from the students' papers to make their points. Carol Miller suggested that one or two less than excellent characteristics should still be a 4. "After all, these are only 10th graders, not professional historians." Carlos disagreed, "how will they know what they are to aspire to in terms of quality unless we hold them to the highest standards? That doesn't mean I wouldn't still give some of the 3s an A, but I don't think we should give the impression that less than outstanding is excellent quality. I'm happy to keep the bar as high as a professional standard." This discussion continued for a long time and a compromise was finally reached. They would expand the point-scale to 5, reserving the highest score for the highest quality work they could imagine -- professional, teacher-generated, outstanding student work. A 4 would be given to an excellent piece of work that covered all of the criteria, but not quite at a professional standard in terms of writing style or reflection. A 3 would indicate very good work, with no more than one flaw in any of the criteria, and not more than three flaws in total. A 2 would indicate several serious flaws in more than one area, and a 1 would indicate a need to substantially redo the piece because very few, if any, of the criteria even approached an acceptable level.

Susan then added some more specific-to-the-task criteria to the original list. As the history teacher who had assigned the work and knew what she was expecting from the students, she wanted to make sure that the students, in addition to all the criteria already decided, were accurate in their information, supported in their interpretations, and that they cited at least three sources besides their oral history interview data.

With this expanded list of criteria, and the agreed upon differentiation between levels of quality performance, the group then tackled the pile of papers that had been originally assigned discrepant scores. As they discussed what, to each, clearly made the paper a 5 or a 4 or a 3, basing their decision on their more refined definitions, they reached consensus rather quickly about systematically applying the criteria. Throughout the discussion they continued to clarify the distinctions between the levels as well as articulate their quality criteria, both by examples and through general descriptions. When one or two of the teachers still disagreed with the majority on a particular item or example, they worked quickly to come to consensus. Their discussion focused on coming to agreement, not on winning the point. They all realized that the importance of this work was in finding criteria and agreeing on how well students met them so that a rather precise assessment protocol could be developed that would serve this class, and, ultimately, the whole school.

With the agreements they'd reached fresh in their minds, the group decided to work in pairs on specific tasks. John and Carlos would draft the language for the rubric that captured the criteria they had decided upon, provide some indicators or examples where necessary, and then work at describing the levels of quality that the group had used to differentiate the rubric categories. They referred to John's sample rubric to see examples of language, get ideas for phrasing, and consider the variations in levels that others had decided to establish. They brought their work to the group in the form of an analytic trait rubric, assigning a score to each of the four traits the group had developed. They said they found that putting all of the criteria into a holistic rubric was too cumbersome, and besides, there were independent criteria that very probably would show different levels of quality. They thought it would be fairer to the students, and more informative to both students and teacher, to assess the independent criteria separately.

Another pair of teachers began to work out the criteria that would be evaluated in the group production of the booklet -- an aesthetically designed compilation of the individual pieces, with introductory passages defining the basic categories of the oral histories. Since the content of the individual pieces would already be assessed, the rubric for the booklet itself, they decided, would be a holistic rubric incorporating aesthetic, organizational, editing, and clarity of presentation criteria. The teachers also suggested to the group that even though the rubric specified 4 levels of quality performance, the final product would have to meet the expectations of a level 4. Since the booklet would be distributed to the public at the museum, it should be no less than an excellent production. The specification of less-than-excellent quality would serve to help students understand what they needed to work on in order to reach excellence.

The remaining two teachers came up with a rubric to evaluate the collaborative effort of the students as they worked together to publish the booklet. Carol and James worked from the list of behaviors that the students used as reminders of successful group-work attitudes and contributions. They decided to use a 3-point scale which defined excellent, acceptable, and unacceptable behavior for each of the criteria on the list. When they returned to the whole group, they recommended that each student complete a self-evaluation of their own contribution to the group using the rubric and that the teachers coaching the group work try to evaluate each student's group participation at least twice during the course of the project. They also suggested that each teacher evaluation be shared with the individual student at the end of the class.

This scenario of constructing rubrics for a complex assessment task highlights a number of recommendations that most advocates of different/alternative assessment practices would endorse.

1. Start with explicit criteria that relate to the outcomes desired from the assignment.
   - If you are formulating criteria for the first time, it is helpful to use a number of resources to guide your decisions. These can come from the task description itself, published rubrics, statements of standards.
   - Read quality student or teacher-generated examples to identify specific characteristics and examples of quality performance.
2. Work with a group of colleagues to help clarify criteria, make them explicit and, ultimately, to share common judgments of quality.
3. Make choices dependent on your instructional and assessment goals regarding the type of rubric you will design (holistic, analytic trait, particular trait), the number of levels of quality you will differentiate among, the kinds of rubrics that will apply to a single complex assessment task, etc.
4. Establish clear distinctions that have instructional relevance between levels of quality.
5. Carefully articulate the various criteria using indicators or examples where feasible.
   - Use examples of quality work to anchor your scoring guide in specific exemplars demonstrating the criteria and the distinctions among levels.

This process of constructing a rubric is particularly relevant for teachers designing one for the first time. Over time, however, this process becomes easier as teachers enlarge their familiarity with formulating important criteria and distinguishing relevant levels of performance. They can then build upon past formulations and achieved consensus. In the case above, the resulting rubric for evaluating the quality of the work was general enough to be used, either partially or wholly, for other assessment tasks. The criteria that define quality organization, style, voice and conventions of written communication can apply to other disciplines as well. Even content criteria, minus the specific references to history content or this particular assignment, can be used as the basis for other task rubrics. In fact, a defining characteristic of good rubrics, according to Wiggins, is that they are "sufficiently generic to relate to general goals beyond an individual performance task, but specific enough to enable useful and sound inferences about the task."[55] As we will see below, using the same or very similar rubrics over time and applying them to a number of assessment tasks has multiple benefits. They help communicate a consistent message to students about essential ingredients of quality performance across a number of disciplines and domains and they also enable schools to reliably chart student performance over time in order to assess how well students are meeting desired educational goals.

The Central Park East Secondary School, for example, has, over the years, developed a single rubric for assessing student portfolios and exhibitions across the curriculum. It is, in this sense, a generic rubric. The language of the rubric has been fully incorporated by teachers and students alike, guiding productively the ongoing teaching-learning exchange.

---

[55] Wiggins, Educative assessment, p. 184.

CENTRAL PARK EAST SECONDARY SCHOOL EVALUATION GUIDE

	VIEWPOINT	CONNECTIONS	EVIDENCE	VOICE	CONVENTIONS	PORTFOLIO ITEM
EXPECTATIONS	Encompasses Wide Knowledge Base and Is Focused: •Clearly identifies, addresses key questions and ideas; •Demonstrates an in-depth understanding of the issues; •Presents position persuasively; •Understands and discusses other views when appropriate.	The Whole is Greater Than the Sum of the Parts: •Introduced and concluded in a satisfying way; •Organized so that all parts support the whole; •Contains useful transitions; •Makes clear the relationships between ideas, notes connections and patterns; •Demonstrates relationships between issues and a larger context.	Credible and Convincing: •Ideas supported by specific, accurate and relevant evidence; •Ideas developed in appropriate depth; •Discusses strengths and weaknesses of evidence where appropriate; •Cites appropriate sources: (primary sources, statistics, quotes from content, graphs, formulas, figures, equations, maps, illustrations) where appropriate; •Analysis of relevant evidence where necessary.	Engaging: •Lively, interesting use of language; •Awareness of reader; •Explains concepts so they are understandable to the reader; •Project has a distinct identity.	Intelligible: •Excellent appearance; •Correct format: (bibliography, footnotes, references, etc. where applicable); •Varied sentence structure; •Good mechanics and standard notation; •Appropriate, broad vocabulary and word usage.	TITLE: STUDENT: ADVISOR: READER:
4	EXCEEDS	EXCEEDS	EXCEEDS	EXCEEDS	EXCEEDS	
3	MEETS	MEETS	MEETS	MEETS	MEETS	
2	APPROACHES	APPROACHES	APPROACHES	APPROACHES	APPROACHES	
1	NEEDS MORE	NEEDS MORE	NEEDS MORE	NEEDS MORE	NEEDS MORE	SCORE: DATE:

COMMENTS: _____

_____

BEST COPY AVAILABLE

Finally, it is important to acknowledge that no rubric is perfect. If you find that the one you have designed does not encompass well the desired outcomes of the task or distinguish clearly enough among levels of quality, it can be adjusted. It is important, however, that the same rubric be used to score all the products of a particular task in order to maintain reliability of the scoring process.

5. Engage students in the evaluation criteria for quality performances

In our discussion of what is wrong with standardized tests, we mentioned the interest that exists in maintaining reliability and validity of test scores through secrecy. Keeping "right answers" secret in order to maintain the "integrity of the test" is completely counterintuitive to good pedagogy. Students are much more likely to produce quality work if they are clearly informed of what is expected of them up-front. In performance assessment practices, the specific criteria used to evaluate student performances are given to students, along with the assignment. The more students understand what is expected of them, see examples of work that correspond to those high expectations and can locate their own work along the continuum of performance quality, the more enabled they are to take responsibility for improving their performance.

This implies some routine practices to encourage high quality work. The first of these is providing students with the rubric that will be used to evaluate their work when it is assigned. If the entire rubric is formulated in language that is unfamiliar to students, they should at least be given the criteria for excellent performance in terms that are accessible to them. If this is a new practice to teachers and students alike, it will probably be necessary to explicitly "teach" the rubric in its component parts. Sharing examples of good writing, good oral presentations, good science experiments, and other performances with students and having them generate the criteria that make them good (in much the same way as the teachers collaboratively came up with evaluation criteria in the scenario above) is a particularly effective way of informing students of the criteria and standards they are expected to demonstrate.[16] It also invests/includes them in the process. Just as the teachers in our scenario found it necessary to anchor their criteria in samples of work that demonstrated the performance progression from novice to expert, students learn to evaluate their own work and progress by seeing examples, with explanations, of work that fits along the rubric continuum. This implies a consistent display of student work that serves as models and benchmarks for students to gauge their own products.[17]

In classrooms where the language surrounding quality criteria is familiar to all, providing helpful, educative feedback to students on work in progress becomes a matter of describing what individual students have done and indicating what constitutes quality performance on any particular trait. Teachers using this strategy have found the work of giving feedback to students much more effective than attempting to edit or correct their products. It keeps the focus on the students and the understanding that they are the primary workers, responsible for moving their work forward. Describing what is there and what should be, in language that is already familiar to students, can, in this regard, reduce the time needed to respond to student work in many teachers' classrooms because the responsibility for revising the work to produce a higher quality product has shifted from the teacher to the student.

Another tool for encouraging students to take upon themselves the responsibility for producing quality work is to construct self-evaluation guides that students apply to their work before they hand in a draft for the teacher to respond to. Pedagogically, we should want students to be reflective about

---

[16] This corresponds to the idea of scaffolding work, making sure students understand the steps along the way to high quality work.

[17] In this regard, it is good to see in many Rural Trust schools walls lined with student work and websites that regularly feature finished student products.

their work, able to assess the quality themselves.[58] As Seidel notes, "...children's perspectives are critical if, ultimately, they are to become their most rigorous and sophisticated critics."[59] We provide two examples of self-evaluation guides below:

Self-Evaluation Check List[60]
Expository Essay

Initial each statement which you feel describes your work. Explain anything you cannot initial, giving reasons. Attach this sheet to your paper.

— My paper has a definite purpose, it 'works.'
— My title indicates my point or slant, and creates interest.
— I had a specific audience in mind as I wrote and revised.
— The details and word choices paint the right picture, give the right feel, and make my point -- without my having to state it.
— I can explain why I placed every punctuation mark as I did.
— Each paragraph contains a clear idea (and topic sentence).
— My opening sentences catch the readers' attention.
— I checked each sentence to insure it was complete.
— I reorganized sentences and paragraphs to make the paper more logical and fluid.
— The transitions are smooth; the paragraphing is appropriate.
— The ending is effective in wrapping things up.
— I read the paper aloud or had others read it to catch awkward, confusing language or mechanical errors.

Project Work Reflections[61]

1. Describe the projects on which you have worked this week.
2. What have you learned while working on the projects? (Give specific examples.)
3. What is the status of the projects on which you have worked this week? What did you finish and what is "in progress"?
4. Classes which involve project work change the way the class is run. More responsibility is placed on the students to manage their time wisely, decide the jobs which need to be

---

[58] When students have developed good self-evaluation, or metacognitive skills, they begin to assume more responsibility for their education. To the question, "how is your science learning coming along?" they stop saying, "I don't know, you can ask my teacher."
[59] Steve Seidel, Joseph Walters, et. al., Portfolio Practices: Thinking Through the Assessment of Children's Work. Washington D.C.: National Education Association, 1997), p. 12.
[60] From Vito Perrone, "Assessment Additions," Harvard Graduate School of Education, 1992. The example comes from the Urban Academy in New York City.
[61] This weekly guide is used by Judy Cross at the Wartburg Central High School (Tennessee), part of the Walden's Ridge Rural Challenge cluster. The guide for work reflections changes from week to week, becoming more comparative -- "compare what you actually accomplished this week with what your intended goals were at the beginning of the week." Judy, a computer teacher, worked with Regina Headden, a teacher of social studies, to design a historical geneology project that reinforced computer technology skills while serving to document the history of each student's family.

completed, set priorities, and meet deadlines. How have you adapted to these responsibilities?

5. Evaluate HONESTLY the quality of your work this week. Think about whether you stayed on task or just "piddled," except when the teacher walked by. How MUCH did you accomplish this week? Are you proud of the work you completed? Would you be proud to have your friends, your family, and your teachers see the work you completed this week?
6. How do you plan to improve the quality of your work, your work habits, your productivity on similar projects in the coming weeks?
7. List specific goals and objectives that you plan to accomplish next week in this class.

Based on what I have described above, my participation grade for this week is:_____.[62]

In relation to the foregoing, we present an outline of a mock trial project in a middle school, humanities/social studies classroom, along with an accompanying self-evaluation form. The latter was constructed by the teacher in a discussion with her students. While the 0-4 scale does not have specified criteria, the teacher regularly discusses the meaning of the scale through exemplars of student work. She asks often, "What would a 4 be like?"[63]

How do You Build a Just Society?
The Trial of Simeone Bushnell and the Fugitive Slave Law
Final Reflections

In this country, we have a complex system for ensuring justice for all. We create laws and then people's actions are judged in courts. Everyone is allowed a lawyer, a judge, and in many cases a full jury. For the past week, we have been enacting this judicial process. As part of your final grade, you must write a thorough and thoughtful reflection discussing the issues of the trial and what you have learned from this project. Use the questions below to guide your thinking. Use a separate piece of lined paper and write in narrative form. You must use complete sentences and paragraphs. Also, be neat and proofread! Mechanics count.

I. Background: *As an introduction, **briefly** explain the case. Be succinct.*

- What was the case about?
- What were the prosecution's main arguments?
- What were the defense's main arguments?
- What do you think of the outcome? In this case, was justice done? Why or why not?

II. Issues:

Simeone committed this crime. We know that he helped John Price escape.

- So, should he be jailed for his crime, or should he go free because he broke an immoral law?
- How do we ensure justice if citizens can take the law into their own hands? And yet, how do we ensure justice if some laws are wrong?
- Even though they knew that Simeone committed this crime, the defense tried to free him on legal technicalities because they believed that the law itself was unjust. Do you think this is right? Why or why not?

III. Conclusion:

- What have you learned from this project about how to create a just society?

[62] Judy Cross told us that students primarily gave themselves high marks the first week she used this Reflection; however, as their work progressed, and as they became more engaged with the project and more familiar with the reflection tool, their self-assessments began to reflect, almost universally, the assessment Judy would have made.
[63] This example was provided by Kathy Greeley, a middle school teacher at the Graham and Parks School in Cambridge, MA.

The Trial of Simeone Bushnell
Evaluation

While everyone wants to win, we are less interested in who "won" and "who lost" this case than in how you "played the game." Use the following criteria to evaluate your own work in our mock trial, based on a 0-4 scale.

1. Preparation

a. Did you know your character and role well? _____
b. Did you prepare for your questions/testimony well? _____
c. Did you do any additional research? Explain. _____
d. Did you have a costume? _____

Comments:

Overall Preparation: _____

2. **Group cooperation:**

a. Did you work cooperatively with others in your group? _____
b. Did you contribute to pretrial planning sessions? _____How? _____
c. Were you encouraging to others? _____
d. Did you use your time well before the trial? _____
e. Did you ever help other people rehearse or prepare? _____

Comments:

Overall Group Cooperation: _____

3. **Performance:**

a. How well did you do in your role? _____
b. Did you stay in character? _____
c. Were you convincing in your part? _____
d. Did you listen carefully when other parts of the trial were happening? _____
e. Were you disciplined and respectful of all parties during the trial? _____

Comments:

Overall Performance: _____

4. Written analysis:

a. Was your final written analysis thorough? _____
b. Was your final written analysis thoughtful? _____
c. Did you write neatly and check your mechanics? _____

Overall Written Analysis: _____

TOTAL SCORE: _____

Any final comments?

6. **Prepare evaluators in order to achieve reliable scoring**

This last criteria for implementing effective performance assessment practices is akin to that of designing rubrics that we presented in the scenario above. The work that the teachers did to arrive at consensus in formulating criteria and anchoring them in exemplars of student work is also necessary to prepare external evaluators (other teachers, parents, community people, field-specific experts) to use the scoring procedures.[64] If the evaluators are not familiar with the criteria being used and how they apply to the work of the students they are evaluating, they will need to practice in order to consistently apply the rubric in the manner intended. Having multiple evaluators place the score they each arrived at on a separate sticky note on the back of the product is, again, a convenient way of comparing scores and scoring practices. In discrepant cases, the evaluators need to achieve consensus and align their evaluations with the interpretation of quality defined by the designers of the rubric. In order for any particular scoring of a single assignment to be reliable, evaluators need only score the work consistently. To be able to compare scores on similar criteria over multiple assignments, however, and to insure that students are applying the same criteria as the evaluators, this consistency needs to be anchored in a common interpretation of quality standards. The best way to insure this is to have evaluators regularly discuss their scoring on actual pieces of work, resolving in consensus any discrepancies that occur. Evaluating products of student performances thus becomes a more explicit, more public practice and one that, when executed with care, will withstand the scrutiny of reliability, validity, and fairness criteria that have traditionally been reserved for standardized, norm-referenced tests, even though the standardized tests don't match the levels we have established for the different/alternative assessment practices.

Portfolios as Assessment Tools

Besides individual performance assessments and exhibitions, which can be related to large projects, public performances, collections of subject matter work, another form that different/alternative assessment can take is that of a portfolio assessment. Portfolios are now commonly found in classrooms and increasingly being required by school systems across the country. They are understood to be a means of seeing student work over time, a way of following growth, of encouraging student self-evaluation.[65] They have gained considerable popularity as an instructional and assessment process.

Portfolios can be compiled in many ways and, depending on what is included, serve many purposes. Some are a collection of all of a student's graded work and performance tasks, samples of "best" work, useful work, most complex work, and the like, as well as tests and quizzes, which are to be passed along to the next year's teacher. Some contain selected samples of student work used primarily to demonstrate to parents or community members the work that students are doing in school. Portfolios can also be assigned to demonstrate the process of performing instructional tasks by showing all the work that goes into a final product. For example, a science lab portfolio might contain all of the paper work that resulted from doing the lab experiment, culminating with a complete written lab report; a writing portfolio might contain a number of pieces at various levels of completion -- from pre-writing explorations, first drafts, revised work to final copy. These selections can show students and their teachers how they managed the process of bringing a task to completion through various stages. It could also help teachers identify strategies or skills that need particular attention and become the data for personal goal setting on the part of the students.

---

[64] These alternative procedures also invite parent participation as students themselves can begin to take the lead in parent conferences, using their portfolios and their self-assessment as a guide. Moving assessment beyond the school opens the possibilities of getting parents and community people involved in the process of assessment, getting closer to the learning process.

[65] It is particularly powerful to see a student portfolio that has been thoughtfully maintained over a year. Student work can be kept track of, the continuities made more clear, the changes in thought and development made more

When portfolios are used for assessment purposes, however, what goes into them and what criteria or standards are to be assessed need to be specified more carefully. Clearly, collecting *all* of a student's work for assessment purposes is unnecessary, not to mention impractical. Whatever is to be considered in a portfolio assessment needs to be articulated in the same way as we have done for a performance assessment task. The standards against which it is to be judged and the levels of competence, from novice to expert, need to be specified in a scoring guide or rubric which is used in order to reliably and validly assess the portfolio. In addition, what goes into the portfolio and who makes that choice is also of critical importance for assessment. If we are to assess, for example, a portfolio of a student's best work, who determines what is "best" -- the student or the teacher? Best from a student's point of view might mean most interesting, easiest to complete, or a host of other factors. Best for teachers, unless established with reference to specific criteria, is also subject to variation. Is it the best effort compared to other students, to this student's prior work, or determined by applying specific criteria? It is usually best for students and teachers to decide together what the portfolio will contain. Portfolios of previously graded work need not be revisited and evaluated again, but an understanding of how the grades were assigned and what they mean needs to be included in order to interpret the work as a whole.

A portfolio assessment, therefore, should be based on criteria other than assigned grades. If grades were the main criteria for assessment, it would only be necessary to consult the grade book. On the other hand, re-evaluating all of the work contained in a portfolio is unnecessarily duplicative. What, then, distinguishes a portfolio assessment from a summation of grades? To answer that question, we need to differentiate the purposes of portfolios as evidence of student achievement. They may be structured in many ways, but for assessment purposes the following are common:

- to show a student's best performances in a number of domains
- to display a student's self assessment of growth and learning, strengths and areas that need attention
- to show the process of learning or the change over time in a student's performance of instructional or assessment tasks.

When portfolios are compiled for assessment purposes, the standards and criteria that follow from the overall goal and purpose of the portfolio are often articulated in descriptions of what it is to contain and how it is to be evaluated. The following examples of such criteria come from Vermont's Portfolio Requirements for Language Arts and Mathematics.

---

evident. Parents typically see their children's work piece by piece. In the process, everything tends to blur,

Vermont Language Arts Portfolio Assessment[66]

Items included:

table of contents	a dated "best piece"
dated letter from the student to the reviewers explaining the choice of the "best piece," as well the process of writing it	a dated personal response to a cultural, media, or sports event; or to a book, current event, math problem, or science concept
a dated poem, short story, play, or personal narration	3 dated prose pieces from any subject area *other than* English or Language Arts

Assessment Criteria

Five dimensions of writing are to be rated on the following levels of performance: extensively, frequently, sometimes, rarely

1. *Purpose* -- Does writer: establish & maintain a clear purpose, demonstrate awareness of audience & task, exhibit clarity of ideas?
2. *Organization* -- Does the writer's response illustrate unity coherence?
3. *Details* -- To what degree are the details appropriate for the writer's purpose and support the writer's main point?
4. *Voice/Tone* -- Does the response reflect personal investment and expression?
5. *Usage, mechanics, grammar* -- Does the writer use correct usage (word choice, tense formation, agreement), mechanics (spelling, capitalization, punctuation), grammar, sentences?

Reviewers were asked to respond to three questions about the school's program (based on reviews of a sample of portfolios):

- Is there progress from earlier dated works to more recent works?
- Is there evidence of variety which will challenge all students and allow for an opportunity for success?
- Is there evidence of teacher/peer response to student drafts, and opportunity for students to revise?

These Vermont criteria statements for assessing portfolios of student work not only indicate the content and criteria for evaluating student performances, they also demonstrate the power of portfolio assessment to analyze and evaluate program design and the quality of the assignments.[67] These criteria enhance the educative value of performance assessment by providing guides for evaluation and, hence, feedback on the quality of instructional activities. If we are concerned with improving students' ability to accomplish valued performances, we need to examine what we are asking students to do by way of assignments and performance assessments as well as how students perform the tasks.

In order for a portfolio to be a useful assessment tool, it needs to be well organized. In the following example of the organizing elements a math portfolio might contain, we see how careful

---

seems to look the same, progress is often hard to see.

[66] Reproduced from Bill Johnson, The Performance Assessment Handbook: Designs from the field and guidelines for the territory ahead. Vol. 1, (Princeton, NJ: Eye on Education, 1996), p. 79.

[67] We have included in Appendix A a writing portfolio from The Peacham School in Vermont that was subjected to the criteria outlined.

BEST COPY AVAILABLE

organization, display and explanation of content allows evaluators to ground their analysis of the included work:

> Organizing elements of a mathematics portfolio might include:
> - a table of contents
> - an introduction or title page that identifies the student and explains what can be found in the collection, and the purpose of the portfolio
> - brief descriptions of selected assignments for readers less familiar with the operation of the classroom
> - labels that distinguish attempted solutions from the final report
> - dates on all entries
> - a review section that includes student reflections and self-assessments, together with teacher comments and peer comments that can help provide important information about the expectations, standards and critical atmosphere in which the various projects were produced.[68]

The kind of thinking and preparation that goes into compiling a portfolio for assessment requires students to reflect on their own learning and to articulate their growth, strengths, and areas still in need of improvement. This kind of reflective, meta-awareness of one's own learning is an extremely desirable product of the assessment process. In addition, not only can a portfolio show growth over time, highpoints of student performances and the quality of assignments, a well constructed portfolio provides an evaluator an opportunity to view the student as a self-conscious learner.

To conclude this section on designing and implementing different/alternative assessments we present a case study of a 9th grade science class that is part of the Rural Trust. Liza Finkel and her students are successfully implementing a variety of different assessment practices connected to place-based project curriculum. There is obviously much to learn from Rural Trust classrooms across the country.

Case Study of a Rural Trust Site Implementing Different/Alternative Assessments
Liza Finkel, Noble High School, Berwick, ME

One of the first things one notices about Noble High School is the cafeteria. Instead of long, institutional lunch tables, students are gathered around small round tables. Lunch time looks more like a restaurant scene than a high school cafeteria. This more personal and friendly atmosphere is mirrored in classrooms and hallways as the 9th to 12th graders go through their day.

We visited Liza Finkel's 9th grade science classroom to learn about the project and place-based curriculum at Noble and the collective work faculty have done around assessing complex, lengthy projects that make up much of the curriculum at the high school. Teachers at Noble have been at this kind of work for several years now and their experience illustrates how place-based, collaborative, real-work curricula have created a challenging and deep educational experience for the approximately 1,000 high school students from three adjacent towns that make up the school population. The communities that the school serves are generally poor. Small industrial plants -- a stadium seating factory, a tannery, an electronics factory and another that makes jet engine parts -- join small businesses like lawn-care and hair-care to provide employment in the towns. Some of the residents are seasonal workers in the tourist industry as Maine's southern coast is only about 10 miles away. Despite a history of people who worked primarily in local shoe and woolen factories, it is now common for Noble high school graduates to pursue higher education. Currently, about 55% of the graduating class enter four-year colleges, another 10% attend two-year institutions and about 8% join the military.

The Rural Trust work at Noble fits seamlessly into the ongoing restructuring that began with membership in the Coalition of Essential Schools about 10 years ago. The organization of the school reflects its implementation of many of the Coalition's Essentials: all students take a rigorous academic program which includes four years of major subjects with all students taking advanced subjects such as physics and math. Classes are heterogeneously grouped: the 9th and 10th grades are organized into teams allowing for interdisciplinary project planning and collaborative assessments. Recognizing the value of the 9th and 10th grade team approach, the 11th and 12th grade faculty are organizing into interdisciplinary houses that will permit more collaborative, collegial experiences for staff and more coherent instruction for students.

---

[68] Seidel, et al. Portfolio Practices, p. 33.

Assessment of student work has been central to the learning that has taken place at Noble high school and can serve as an example to other Rural Trust sites embarking on complex curriculum projects that connect students and community members to place. Lisa's science class provides a good window from which to view their work.

The 9th grade science curriculum is entirely project-based. Students work through six large projects that are comprised of a number of individual and group tasks and products. Each project is outlined for students at the beginning, providing them with the Essential Questions that will guide their inquiry into the subject, a list of Goals or intended outcomes, and how the project meets state mandated Learning Results, Maine's term for standards. The first project of the year, the Stream Project, pursues answers to these Essential Questions:

1. *How do people study, explain, and represent the world using scientific tools?*
2. *What can a stream tell us about the quality of our local environment?*
3. *How can I make a difference?*

The project examines the role of water in our lives and in the natural world. Students use the natural laboratory of the local Worster Brook to carry out a series of assignments which include the following:

Major Assignments for the Stream Project

- **Water Cycle Diagram**
  In this assignment, you will develop a more complex view of the water cycle than the one you may already have. Through participation in a simulation game, you will travel as a water molecule might through the water cycle. You will then be asked to represent your understanding of this more complex water cycle with a detailed diagram or written explanation as well as to reflect on what was new about the content of this exercise and what was review.

- **Groundwater Exploration**
  In this assignment, you will explore in more detail the part of the water cycle during which water is underground. As a class, we will investigate what happens to water as it passes through groundwater systems and use that knowledge to design and evaluate water purification systems.

- **Visual Survey and Habitat Inventory of Worster Brook**
  In this assignment, you will work with a partner to begin to describe Worster Brook and the environment that surrounds it. Using a checklist designed by the Environmental Protection Agency (EPA), you will gather data on the physical characteristics of Worster Brook and will prepare a written report on your findings. Your report will end with a list of questions you and your partner are interested in investigating further.

- **Stream Guide Contribution**
  In this assignment, you will work with a group of students to collect and identify plants, macro-invertebrates, and other organisms from the stream and associated riparian zone. As a class, we will calculate the Pollution Tolerance Index for the stream to get a measure of stream water quality. You will then create a "Stream Guide Page" to describe one organism's habitat, niche, life cycle, and pollution tolerance.

- **Chemical Analysis of Water in Worster Brook**
  In this assignment, you will collect and analyze samples of water collected at the brook. You will learn how to systematically collect water samples, and how to conduct chemical tests to determine the amount of nitrates, phosphates, and dissolved oxygen present in brook water samples. You will also test for the presence of fecal coliform and determine the pH of the water. You will also learn how to interpret your results and to determine what your tests tell us about the quality of the water in Worster Brook.

- **Freshwater Ecosystem Concept Map or Diagram**
  In this assignment, you will be asked to make a concept map or labeled diagram illustrating the ways in which all of the concepts we have been studying are linked to one another. This diagram should also show the ways in which we have seen examples of these concepts in our study of the stream ecosystem.

- **Design, Write-up and Presentation of Stream Investigation**
  In this assignment, you will design and conduct an investigation into at least one aspect of the stream that you are interested in learning more about. Your investigation should demonstrate what you have already learned about the stream through our work as a class. You will share the results of your investigation with the rest of the class at a "Share Fair" in which you will display your findings and summarize your results in writing and out loud.

- **Community Contribution**
  In this assignment, we, as a class, will decide how we can best communicate what we have learned about Worster Brook to the community.

- **Stream Project Portfolio**
  For this assignment, you will gather all of the work you have completed for the Stream Project and organize it into a portfolio. Keeping copies of your work in a folder as you go along will make this assignment very easy! In addition to compiling the work you have already completed, you will create a Table of Contents, an Introduction and a Conclusion.

BEST COPY AVAILABLE

These major assignments, which are worked on over the first eight weeks of school, are composed of a variety of instructional activities that thoroughly familiarize students with all aspects of water. "In fact," Liza remarked, "we go into it so thoroughly that students tell me I ruin water for them." They spend 80 minutes collecting macro-invertebrates, indicators of water quality, and then 80 minutes the next day categorizing them. All in all, students visit the Worster Brook at least four times over the course of the project -- once to conduct a visual survey, once to collect macro-invertebrates, once to collect water samples for chemical analysis, and at least once to carry out the scientific investigations they have designed on their own in groups. They use EPA materials for visual and site habitat monitoring. Students calculate how much water they use over the weekend and they read local newspaper articles about the state of their local water conditions. They learn to conduct chemical analyses and each student becomes an expert with a particular test. Ultimately, groups are comprised of appropriate experts to work on a particular scenario of water pollution. Based on their scenario, they predict what test results they would expect and what the community should do to fix the situation.

The assessment of the work is ongoing as Liza consults with groups of students while they work on their assignments, but it is regulated by a detailed rubric which outlines not only what is required for each major assignment but describes levels of quality performances. Noble High School uses a 5 level scale for most of its rubrics -- Distinguished, Advanced, Proficient, Novice and No Credit -- which correspond to the 5 letter grades A-F. Liza only describes what goes into the Distinguished, Advanced, and Proficient categories because she requires all of her students to aim for at least a proficient performance. Assignments in Liza's class are always available to revise or to add to as students understand better what is needed to produce a high quality product. Liza also relies on last year's assignments as models for this year's students. She says that seeing what students have done in the past is often more meaningful than reading directions and students strive to improve the quality of assignments from the previous year. It becomes, for them, a challenge.

The first rubric students see in the year is given to them as they revise their Water Cycle Diagrams. In this way, Liza teaches them to use the rubric as a quality management, revision tool and introduces them to the value of rubrics, which they will use for each major assignment from that point on.

RUBRIC FOR REVISED WATER CYCLE DIAGRAMS
9th GRADE SCIENCE FALL 1998

This rubric will give you some guidelines as you revise your water cycle diagram. We will use it to grade your final draft which is due on Thursday, October 1. Be sure to turn in the rubric and the first version of your diagram along with your final draft.

PROFICIENT
Includes all seven "stations" from the water cycle simulation: stream, animal (human and other), plant, mountain, ocean, groundwater, cloud

Shows connections (using arrows) among all stations (remember, there is more than one way to get to the cloud, as there is to many of the other stations)

Labels at least two connections from one station to another with a descriptive term explaining HOW the water travels from one place to another

Is legible and neatly done

Shows effort; (was not done in five minutes before class started)

ADVANCED
Meets all requirements for a Proficient grade

AND

Includes other places water might travel (in addition to the seven listed above)

Labels ALL connections from one station to another with a descriptive term explaining HOW the water travels from one place to another

Is creative (uses color or illustrations to add interest and information to the diagram)

DISTINGUISHED
Meets all requirements for an Advanced grade

AND

Includes details about what happens to water when it is in each station (for example, what happens to water when it is underground or in the ocean)

Shows connections between the water cycle and YOUR life

Includes information about sources and effects of water pollution

As the year progresses, students become so familiar with using a rubric that "they ask for it even before I give it out," says Liza. "They get used to it and expect it." She sees a tremendous value in letting all students know what they need to do to get "Distinguished," and overhears students talking about what kind of evaluation they are aiming for. Knowing the standards helps her students with self- evaluation, which figures in the portfolio assignments in her class as well as those produced each semester by every student in the school.

Another assignment Liza gives the students during the Stream Project engages the students in real work -- designing a water purification system. The design of the facility is rated by the entire class for efficiency based on timed trials performed in class and a comparison of water samples from the various designs. The task description for this authentic assessment indicates the roles that students assume as they complete and assess the work.

Name: _____ *Section* _____

HELP WANTED: Water treatment facility designer

Noble High School needs to build its own water treatment plant in order to meet the needs of its growing student body. Each group in this class is being given a chance to submit a design proposal for the new facility. Each group will have equal opportunity to come up with its best design using the materials provided by the facilities planning committee. The facilities planning committee will then review the proposals and evaluate how well they have met the following criteria:

- *The facility must work as **quickly** as possible*
- *The facility must be as **effective** as possible*
- *The facility must be as **inexpensive** as possible*

*The easiest way to convince the committee that your group has come up with a good design is by keeping good records of your materials, procedure, and results for each trial. Fill out the "trial records" provided on this sheet as you test each design. **(You MUST fill these out!!)** These records will support your claim that yours is, in fact, the best water filter design.*

Trial #1

Materials used

Procedure

Start Time: _____
Stop Time: _____
Time Lapse: _____
How effective was this filter? (Very, Moderately, Not At All) _____

The Stream Project is the first of six that make up Liza's curriculum for her 9th grade science class. Some are interdisciplinary -- like the joint English, history and science project on Exploration (the Essential Question guiding the project is: Can there be exploration without exploitation? through which students look at European exploration of the "New" World in social studies and study Mars exploration in science. While reading utopian novels in English they prepare themselves to design a non-exploitative colony to explore Mars). Another interdisciplinary project results in a mock UN Summit to identify the 10 most pressing world issues to be addressed in the next decade. This culminating project in the Freshman year, that combines math, English, science and social studies and to which townspeople are invited, is an all day event held in a local church. Liza's other projects explore more conventional science topics such as plate tectonics, scientific thinking and energy.

Each project concludes with a portfolio of the assessment pieces of the assignment, a table of contents, an introduction, and a conclusion. The latter are reflections on students' learning over the course of the project. Reading the introductory and concluding narratives of each of the portfolios lets Liza know what her students learned.

> *[The portfolio] informs my teaching a lot by asking an open-ended question about what they think was the most important thing they learned in the unit. It's much more informative than any unit test I could give them. And they need to learn how to talk about their work. It is hard, though, because some kids don't like to write very much. The introduction and conclusion pieces are big for them, though only five out of 80 kids didn't turn one in in the latest project.*

The following example of a project portfolio guide, along with a related assessment rubric, is the first one of the year and culminates the Stream Project. (With some modest reworking, it could be used for units in other fields or in interdisciplinary work.)

Stream Project Portfolio Guide
Fall 1998

What is a Project Portfolio? A project portfolio is a collection of the work you have done on this project. You MUST include certain pieces of work (see list below) and you can CHOOSE to include others (see suggestions below). This portfolio should demonstrate what you have learned through doing the project, and should also allow you to show off your best work.

What MUST be included in your portfolio:

- *A Table of Contents page*
- *An Introduction in which you introduce the project as a whole, summarizing or listing the essential questions, and briefly describing the purpose of the project*
- *Water Cycle Diagram*
- *Stream Guide Contribution*
- *Design and Write-up of Stream Investigation*
- *What do you know about water quality? assignment*
- *At least one homework assignment (of your choice)*
- *A Conclusion in which you:*

(1) Summarize our findings about the water quality of Worster Brook, including information from the Visual Survey and Habitat Inventory, Macro-invertebrate Study, and Chemical Tests.

(2) Explain how the work in your portfolio demonstrates what you have learned through this project (refer to the ME Learning Results on the initial handout for this).

(3) Answer the following questions:

(a) What is the piece of work that makes you the proudest in this portfolio? What about it makes you feel that way?

(b) What is the piece of work you wish you had done a better job completing? What would you do to improve it if you had a chance, to revise it further?

What else MIGHT you CHOOSE to include?

Copies of other assignments you have completed for this project
Copies of related journal entries
Work that you have done on your own, or for another class, related to this topic

BEST COPY AVAILABLE

RUBRIC FOR STREAM PROJECT PORTFOLIO

A Proficient Portfolio

- *Contains all required assignments*
- *All assignments are complete and revised to at least the Proficient level*
- *Includes a legible Table of Contents*
- *Includes a clear and complete Introduction*
- *Includes a clear and complete Conclusion*
- *All pieces (including Table of Contents, Introduction and Conclusion) are either typed or legible*
- *Is neatly organized (at a minimum, assignments are placed in the portfolio in the order listed on the Table of Contents)*

An Advanced Portfolio

- *Meets all of the requirements for a Proficient Portfolio*
- *All assignments have been revised or corrected to the Advanced level*
- *Includes at least one optional assignment*
- *The Introduction and Conclusion are detailed and show evidence of thought and effort beyond the minimum requirements*
- *Is presented neatly in a folder or binder*

A Distinguished Portfolio

- *Meets all of the requirements for an Advanced Portfolio*
- *Includes at least two optional assignments*
- *Most assignments have been revised to the Distinguished level, and any that are not Distinguished are revised to the Advanced level*
- *Demonstrates connections between the work of the project and related ideas or issues from other classes, or from outside of school (these connections may be included in the Introduction and Conclusion of the Portfolio, or may be included as extra assignments)*
- *Demonstrates the use of outside resources*
- *Is unique and distinctly the work of you as an individual*

Learning to be organized is one of the biggest challenges for her Freshmen students as they compile their portfolios. Students' reactions to this task vary, according to Liza. "Sometimes they love it and sometimes they hate it. It is a huge organizational task -- especially early on in the Freshman year. They have to remember to save stuff." But she also noted that students become adept at the process during the year. She is able to hold up a portfolio later on in the semester and students will point out that there is no table of contents or no introduction. "They notice what is missing. They number pages now." Some of her students devised what Liza thought was a rather ingenious way of marking their place for specific pages -- "they designed tabs out of masking tape. It's hard to tell you just how phenomenal that is for this group of boys."

The final assessment in Liza's class is a portfolio that requires students to collect the project work they have done throughout the year and write a self-reflective introduction and conclusion. The task description for this portfolio illustrates the kinds of open-ended questions Liza asks throughout the year to encourage students to evaluate their own learning.

Science Final Experience
Spring, 1999

This final experience is designed to help you reflect on your past year in science.

In order to do this, you will complete three activities. Your overall grade will be based on:

- *The thoughtfulness of your participation DURING the final experience on EACH of the tasks*
- *The quality of your individual contributions today (written and oral)*
- *Your ability to demonstrate connections between activities or units from the past year and the activities you complete today*
- *Your ability to contribute and work with others in small and large groups today*

Part One: Written Reflection (on a separate piece of paper)
In this part of your Final Experience, you will be asked to answer, in writing, the following reflection questions. You will be graded on the thoughtfulness of your answers, and on the quality of your answers. This means that you should write in complete sentences, be specific, include as many details as you can, and reflect carefully on what you have learned over the course of the semester and the year in each of your answers.

1) A. What would you tell a friend from another team, or from another school, about the effects our every day activities have on the earth's ecosystems, and what we should do about them?.
B. How could you convince them of the seriousness of this issue?

2) A. What was the most useful, interesting, or important thing you learned in science this year?

B. What made it that way?

3) A. What was the piece of work you completed this year in science that makes you the proudest?

B. Why does it make you feel that way?

4) Looking back over your freshman year in science, what would you do differently if you had the chance to do it all over again? Why?

5) What advice would you give next year's freshmen to help them be successful at Noble High School?

Part Two: Letter to a Future Teacher (on a separate piece of paper)
In this part of your Final Experience you will be asked to write a letter to a new teacher. Your letter should include your idea of what makes someone a good teacher, and advice to a new teacher that will help her or him be the kind of teacher who can help someone like YOU learn. This letter will be given to a real person, and I have given you their name below so that you can make it as personal as possible. Be sure to tell them something about you (your grade level, your interests, your name, etc.) as well as give them good advice that they can use as they begin teaching. You should proofread your letter and turn in a FINAL DRAFT (no spelling errors, no crossed out words, etc.).

You should address your letter to: _____

Part Three:

A. Saving the World's Resources
In this task you will be asked to solve a problem as a whole class. You will be given a limited amount of time to decide on and then carry out your solution together. You will be graded individually on your participation in this activity, and as a group on your ability to work together to solve the problem.

B. Reflection and Discussion
You will be asked to reflect as a part of a class discussion, on the activity you have just completed. The following questions should guide the discussion:

1) What would you describe as the strengths of the LARGE group in solving this problem? (what did you do that worked)?
2) What were your individual contributions to solving this problem?
3) What did you learn from solving this problem about the group, and about yourself?

Not all teachers at Noble use portfolios in lieu of unit tests as Liza does, but the whole school requires each student to compile a portfolio of their "best work" each semester. The portfolio is continually updated during their high school career and the Senior Celebration includes a public presentation of it. To ensure that students keep their work over the course of their four years at Noble, their homeroom teacher keeps the portfolios in a file

cabinet throughout the year and over the summer. Students know where to go for their portfolio when it is time to update it and pass it in to their current homeroom teacher.

The teachers at Noble also have to contend with the standardized tests that are required in Maine. In general, the MEA (Maine Educational Assessment) scores have been good. They have gone up in every area but math and, overall, the scores have improved since the beginning of restructuring with the Coalition of Essential Schools. The biggest risk they face is the fact that the MEA is now a 9-day marathon test. As Liza notes, "Because our kids are used to adults treating them as adults, taking responsibility for their actions, becoming independent thinkers, sitting for 9 days to take tests is not something they want to do." There is some worry that students will not perform well on these accountability measures as a result. Also, since all students are required to take physics at Noble in their senior year, but the MEA tests for physics in the 11th grade, there is a concern that students will not demonstrate the knowledge that most will actually acquire after the standardized test is administered. So far, the MEA is not a high stakes test -- a requirement for promotion or graduation. They inform students of quality criteria for each assignment and engage them in self-assessment throughout their high school years. In the meantime, however, the staff at Noble has found that a curriculum based on projects that connects students to their local environment as well as to serious issues facing their students as members of their community, citizens of the state of Maine and of the world has not diminished the scores on generalized achievement tests such as the MEA. Quite the contrary. Liza notes that since the introduction of project-based learning which is connected to their real life/locale, heterogeneous grouping, and a rigorous academic curriculum for all students, "everything has gone up. We have raised the aspirations of the kids and the community."

From her experience in designing and implementing projects and different alternative assessments, Liza offers these helpful suggestions for communities and teachers beginning this work:

1. *Provide support for teachers -- especially time, resources, and relief from curriculum coverage pressures -- to help them design and develop this kind of curriculum.*

2. *Get examples of other people's curriculum, preferably work that is not in one's own discipline, in order to get ideas of how to construct projects and assessments.*

3. *Be willing to try. If it doesn't work the first time, try to think about what is working and why. Listen to the students and see what they think.*

4. *Know what you want students to learn, what matters most (this is the most basic part of the rubric), and what you want them to do. You need to know this well enough to be able to write it down. If you cannot, you are not ready yet.*

5. *Using central, essential questions gives you the big picture and specific goals for individual assignments help orient students in their attempts to answer the larger questions.*

6. *Rubrics help students take the mystery out of quality work. Seeing examples of work from previous years raises the quality of student work because they experience directly what good work looks like.*

How Can We Observe Student Learning with Performance Assessments?

Whether assessment is conceived of as an audit of student knowledge or memory, as some try to do with standardized or traditional teacher tests, or as a means of educating students about quality performances that are valued in the world beyond school, as in the different assessment practices we have described here, the question that teachers and parents want answered is: How do we assess student growth in learning? The simple answer is: by observing it. As we have already stated in the discussion of standardized testing, one of the major deficiencies that results from emphasizing test scores is the inability of those tests to provide valid information about a particular student's knowledge, let alone know-how. The results of standardized tests are provided in numbers and percentiles that say very little about what a student can do, what he or she needs to work on, or what specific changes in

BEST COPY AVAILABLE

learning the student has achieved. These results lack the information and the timeliness to provide an appropriate response to a student's learning needs.[69]

Making use of rigorous performance and portfolio assessments provides opportunities to observe student learning over time and give educative feedback to students and their teachers in order to enhance students' ability to learn valued performances. If there is general agreement on what constitutes an education of value, specifically articulated instructional goals, established, agreed-upon criteria for distinguishing among levels of quality of the performance, a careful, reliable scoring of student work using common criteria, and a system of documenting that scoring, in its specifics, over time, these different/alternative assessment practices can provide a frame for observing student learning. And observation of learning is the first step to responsively impacting it.

Let's imagine a school that has developed scoring guides for the agreed-upon and valued performances that its students should demonstrate. There are many general rubrics for different kinds of writing, oral presentations, science lab experiments, mathematical problem-solving and reasoning, research reports, and collaborative and independent work habits. These general rubrics are modified according to the specific assessment tasks assigned to students. Student work, across various disciplines, is scored by traits that are common to the performances being assessed. Both within a particular discipline or class as well as across disciplines, student work can be examined over time to see if and where there is evidence of growth and where there is need for improvement. This information can be quite specifically geared to individual students if the scoring of work across similar dimensions is consistent and rigorous and, importantly, includes conversations among teachers about those dimensions as they pertain to particular students. One way of keeping track of student growth over time, as we have seen above, is to collect samples of scored work in portfolios for assessment of the learning taking place. Another way to observe learning over time or trends in class performances (an indicator of program effectiveness) is to maintain a data-base of scores on individual traits over multiple assignments. This record-keeping of student performance data is one way of observing if there are variations in performances, what those variations are (positive, negative, or random), and what achievement or learning can be implied by these scores over time. Of course, a true picture of a student's learning/ability can only be observed by carefully examining the progress demonstrated in actual pieces of student work, but keeping track of carefully scored assessment tasks for individual students can be presented in such a way as to focus the examination of pieces of student work on specific traits or questions.

A simple example of keeping track of sample data of student growth in writing based on two different writing assessments using similar rubrics can be seen in the sample data below (abbreviated for this presentation):

---

[69] One of the authors was recently at a school planning meeting of teachers, administrators and parents in which they discussed the elements of a new literacy program that will be required for all entering 9th graders who score "poorly" on the district required standardized test. Given the testing company's time frame for communicating scores, however, the students may be placed in the 9th grade program on the basis of their 7th grade scores. Not only is the test inadequate for assessing student learning, this use of it assumes that students <u>don't</u> learn over time.

NAME OF STUDENT	1 C&O	1 S&V	1 M	1 COMMENTS	2 C&O	2 S&V	2 M	2 COMMENTS	1–2 C&O + / –	1–2 S&V + / –	1–2 M + / –
Petunia, Sally	3	2	3	Repetitive; weak vocab.	3	3	4		0	1	1
De Pina, Daniel	2	2	3	Missing para, poor intro	2	3	3	No thesis, weak argument	0	1	0
Gainer, Tamara	3	3	3		3	4	5		0	1	2
Jones, Melinda	2	2	0	Work on spelling and punctuation	2	2	1	Weak argument	0	0	1
Deal, Jason	3	3	3		3	3	4		0	0	1
Ortega, Anne Marie	2	3	2	Unclear ideas, no thesis	2	3	2		0	0	0
Dowling, Chauntee	3	3	4		4	4	3		1	1	-1
Williams, Keshia	2	2	3	Short sentences, no variation	3	2	3	Simple sentences, repeats	1	0	0
Moore, Janeida	3	3	4		4	4	5		1	1	1
Story, Corine	3	3	4		4	4	4		1	1	0
Hung, Tran	3	3	3		3	3	4		0	0	1
Holmes, Jada	2	3	3	No conclusion, weak argument	3	2	2	Poor word choice, spelling needs work	1	-1	-1
Jones, Paul	3	2	3		3	3	5		0	1	2
Amado, Alfie	3	2	3		3	3	2		0	1	-1
Mui, Cindy	3	3	4		4	4	4		1	1	0
Duchane, Rene-Yves	2	3	3		3	3	4		1	0	1
Lewis, Lakesha	1	1	3	No general ideas, only examples	3	2	2	Work on punctuation	2	1	-1
Claiborne, Chris	2	2	2		3	3	2		1	1	0

Legend:
C&O = Content and Organization
S&V = Style and Voice
M = Mechanics (spelling, grammar, punctuation, formatting)
1 = 1st Writing assessment
2 = 2nd Writing assessment
1–2
+ / – = Change in score from 1st to 2nd assessment by trait

These raw data, particularly the comments, can give information to the teacher and students about changes, improvements and particular needs in writing demonstrated on individual assessments and by comparing similar assessments for the same students. The final three columns chart the change in scores on the three traits used for scoring the writing. Not only does this grid give information on individual students, it also provides feedback to the teacher about the success of her instruction for the group of students as well as for individuals. One might infer from these data, for example, that students showed the most growth in mechanics and then in style and voice. If this was the intended focus of instruction, this way of presenting data could generally confirm the instructional strategies used. If, however, the teacher's intentions were to improve the content, ideas, and general organization of student writing, these data might indicate a need to re-examine instructional strategies and goals.

These types of data are also helpful for tracking student learning in particular areas. Any number of ways of displaying these data can be developed from simple data bases of scores. The variety of items that can be identified for examination can make visible important information about student performances from which inferences about learning and teaching can be drawn.

We include two examples of charted data to show the variety of lenses on student performances that an analytical scoring of assessments can provide.[70] If we want to know, for example, the number of students who scored at the highest levels on a particular assessment to compare to subsequent assessments or perhaps to the performance of students last year, we might want to display the data similarly to the following graph.

[70] The examples are taken from Herman, A practical guide to assessment.

BEST COPY AVAILABLE

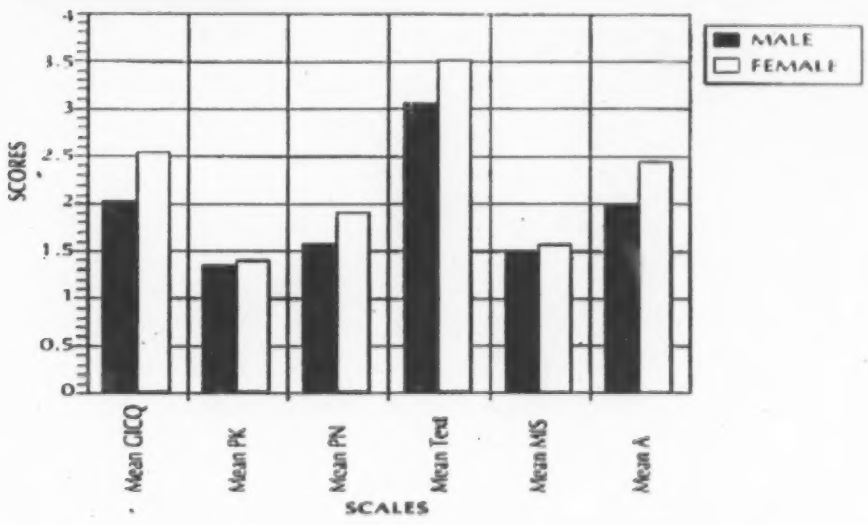
Distribution of Student Scores
Biology Laboratory Procedures Assessment

NUMBER OF STUDENTS								
	1	0	1	1	3	5	6	5
							X	X
						X	X	
						X	X	X
					X	X	X	X
					X	X	X	X
	X		X	X	X	X	X	X
	1	2	3	4	5	6	7	8
X = Each student receiving a particular rating from 1-8								

We learn from the chart that most students (16 of 22) are at the three highest levels. We could ask, how do this year's scores compare to scores last year? or how do these scores compare to other biology tasks students have completed?

We might also have an interest in the performance of boys and girls on various tasks. The graph that follows compares their respective performance on their history essays in relation to a set of criteria.

Profiles of History Essay Performance for Boys and Girls

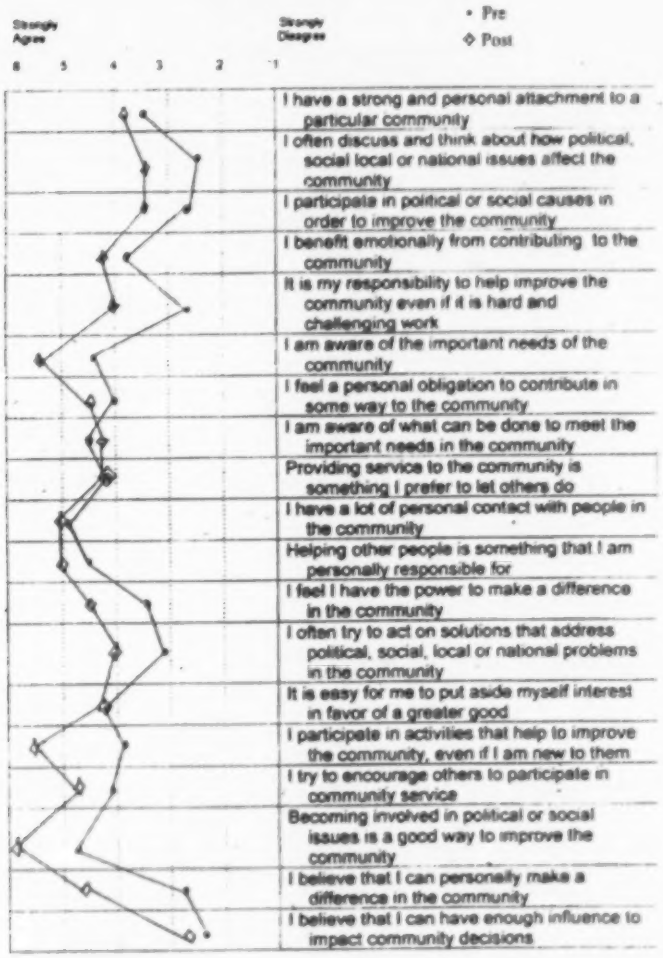


ESSAY DIMENSIONS: GICQ=General Impression of content quality; PN=Number of Principles or Concepts; PK=Prior Knowledge; A=Argumentation; TEXT=Text (proportion of essay using text-based detail); MIS=Misconceptions

The graphs above are only samples of how information can be displayed in graphic representations to indicate any number of dimensions of student work. Whether we wish to ask questions about an individual performance, one student's growth in several domains, compare performances of the entire class or this year's class with last year's, and so on, we can use the scores that have been assigned to assessments in a rigorous and consistent manner to provide us with an image of trends. By closely examining student work in light of the trends made visible by creative displays of the scoring data, we can see even more deeply into the learning, growth and understanding that students demonstrate on well conceived, complex, stimulating, real-world valued performances. Moreover, some of the scoring data in graphic form can be presented as part of a school's public accounting to the local community, leading to discussion about ways of addressing issues that seem in need of attention.

While the following graph is not linked to performance assessments *per se*, it is another demonstration of how schools can tap into other areas of student growth and track it over time. This graph is part of the work that the North Coast Rural Challenge Network produced for its 1997-98 evaluation report. It provides a way of charting changes on a survey used to assess students' awareness of community concerns and their commitment to taking action for the benefit of the community. The North Coast Rural Challenge Network gave the survey before students engaged in community-oriented project work and then again at the end of the year. Their goal was to begin to measure and keep track of desired changes in students' attitudes and behaviors as a result of their Rural Challenge work.

Rural Challenge Survey Graph


Strongly Agree
Strongly Disagree
• Pre
◇ Post
6
5
4
3
2
1
I have a strong and personal attachment to a particular community
I often discuss and think about how political, social local or national issues affect the community
I participate in political or social causes in order to improve the community
I benefit emotionally from contributing to the community
It is my responsibility to help improve the community even if it is hard and challenging work
I am aware of the important needs of the community
I feel a personal obligation to contribute in some way to the community
I am aware of what can be done to meet the important needs in the community
Providing service to the community is something I prefer to let others do
I have a lot of personal contact with people in the community
Helping other people is something that I am personally responsible for
I feel I have the power to make a difference in the community
I often try to act on solutions that address political, social, local or national problems in the community
It is easy for me to put aside myself interest in favor of a greater good
I participate in activities that help to improve the community, even if I am new to them
I try to encourage others to participate in community service
Becoming involved in political or social issues is a good way to improve the community
I believe that I can personally make a difference in the community
I believe that I can have enough influence to impact community decisions

BEST COPY AVAILABLE

Public Accounting

As noted, data from the different/alternative assessment practices can be organized for purposes of reporting to external audiences, fulfilling accountability requirements. We believe, in fact, that Rural Trust schools should make their work and the work of their students public. The New Country School (in Minnesota and part of the Center for School Change) holds public meetings every eight weeks at which students present their work. We can imagine holding regular (possibly twice a year) science fairs, history fairs, math fairs, language and literature fairs, and the like, where students would present their portfolios for review, where student projects could be examined for their quality. And we believe annual public assessment reports should be issued by the schools which bring together data from the various assessments discussed in this monograph. These public reports can be responded to by parents and community people and can be the basis for a process of School Quality Reviews conducted by teams made up of teachers, administrators, parents and community people (along with a higher education representative and a representative from the State Department of Education).

Assessment rooted mostly in standardized tests means essentially that very little accountability exists in regard to genuine standards and concerns for high quality education. We can do better than this.

Closing Statement

We began this monograph with an explanation and critique of standardized tests. Our fundamental criticism of them is that they provide meagre information concerning complex understanding and know-how, qualities that are surely relevant in the real world beyond school. We saw how the overriding concern with comparing students in order to discern the "best" and the "worst" -- which is built into the bell-curve distribution of scores that are norm referenced -- actually prevents most standardized tests from providing any educationally useable information about students' understandings. The practices implied in rigorous, complex, reliably evaluated, and carefully documented different/alternative assessments provide the possibility of looking more deeply and understanding much better how students construct valued performances. Further, it is the basis for serious self evaluation on the part of students. Creative displaying of the data different/alternative assessments yield can indicate trends, nuances of performances and growth and learning that can be examined in their complexity by carefully reviewing examples of student performances over time with particular questions or inferences in mind.

Although it clearly takes more time, more collaborative support from colleagues, parents and community evaluators, the implementation of different assessment practices seems well worth the effort. The information about learning and teaching that they can reveal sets the groundwork for creating, modifying and tailoring educational practices to meet the particular goals and desires of communities and their schools. This is what genuine accountability is about.

Bibliography

Burke, K., R. Fogarty & S. Belgrad. (1994). The mindful school: How to assess authentic learning. Arlington Heights, IL: IRI/SkyLight Training and Publishing, Inc.

Darling-Hammond, L., J. Ancess & B. Falk. (1995). Authentic assessment in action: Studies of schools and students at work. New York: Teachers College Press.

Gould, S. J. (1981). The mismeasure of man. New York: Norton.

Greenberg, K., H. Weiner, R. A. Donovan. (1986). Writing assessment: Issues and strategies. White Plains, NY: Longman.

Herman, J. L., P. R. Aschbacher, et al. (1992). A practical guide to alternative assessment. Alexandria, VA: Association for Supervision and Curriculum Development.

Johnson, B. (1996). The performance assessment handbook: Designs from the field and guidelines for the territory ahead. Vol.s 1 and 2. Princeton, NJ: Eye on Education.

McDonald, J. P., S. Smith, et. al. (1993). Graduation by exhibition: Assessing genuine achievement. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. J., J. S. Kendall, et al. (1999). Essential knowledge: The debate over what American students should know. Aurora, CO: Mid-continent Research for Education and Learning.

Meier, D. (1974). Reading failure and the tests. New York, Workshop Center, City College of New York.

Newmann, F. M., G. Lopez, et al. (1998). The Quality of Intellectual Work in Chicago Schools: A Baseline Report. Chicago, IL, Chicago Annenberg Research Project.

Perrone, V. (Ed.). (1991). Expanding student assessment. Alexandria, VA: Association for Supervision and Curriculum Development

Perrone, V. (1992). Assessment additions. Cambridge, MA: Harvard Graduate School of Education.

Seidel, S., J. Walters, et al. (1997). Portfolio practices: Thinking through the assessment of children's work. Washington, D.C.: National Education Association.

Spandel, V., R. J. Stiggins, (1997). Creating writers: Linking writing assessment and instruction. Second Edition. White Plains, NY: Longman.

Wiggins, G. (1993). Assessing student performance: Exploring the purpose and limits of testing. San Francisco: Jossey-Bass.

Wiggins, G. (1998). Educative assessment: Designing assessments to inform and improve student performance. San Francisco: Jossey -Bass.

Websites

www.classnj.org The Center on Learning, Assessment and School Structure provides a cooperative web page which publishes a number of rubrics, assessments and teacher-developed curriculum. The site is developed by Grant Wiggins and his associates.

www.aisr.brown.edu is the website of the Annenberg Institute for School Reform at Brown University. The site offers information and practical examples of assessments and protocols for teachers looking at student work together.

www.mcrel.org is the Mid-continent Research for Education and Learning site listing useful information and resources for assessment practices.

Appendix

We include here a sample of a 5th grade writing portfolio provided by Kathy Renfrew and Margaret MacLean from Peacham, Vermont. As mentioned in the text, Vermont requires portfolios of student work in writing and mathematics to be scored by the state. The one included here was prepared for the state, which randomly selects portfolios from each school in the state at the 5th and 8th grade levels, but was not selected for evaluation. Instead, as with all of the portfolios that are not selected, the classroom teacher and another teacher in the school score each piece of writing and provide feedback to the student in the form of scores and some comments. Teachers who evaluate portfolios are trained to use the very elaborate and detailed rubrics that are being used by the state of Vermont in the assessment process. We include the student work that comprised this portfolio. According to the guidelines, the student has selected from among her large collection of writing for the year.

Students at Peacham Elementary School keep portfolios of all of their written work for each school year. This includes all preliminary and final drafts of the work of the student. In the spring of the year, students select their best piece of writing for each of the genres required by the state (writing to instruct, report writing, persuasive writing, narrative writing, response to literature, reflective essay). This selection is made in consultation with the teacher.

Even though portfolio assessment at this level attempts to move well beyond the limitations we noted for standardized tests, undertaking this at the state level can lead to some of the same problems we noted in the case of large scale standardized tests. For example, students only receive their state-scored portfolio at the beginning of the next school year, and then the evaluation is only a score for each product and an overall score for the portfolio. No comments that might inform the students of what they are demonstrating in their writing or what they might need to pay more attention to are provided. This makes the work of the classroom teacher much more valuable in the day-to-day monitoring of writing performances, as well as providing detailed feedback on numerous drafts and determining a rubric score for the final draft. Communities of readers of student writing in a local area are more likely to provide timely and useful response to the work.

It might be interesting for a group of teachers and parents to read the work of this fifth grade child and discuss and rate the writing in relation to one of the writing rubrics we have presented previously in the text.

I AM A MEDIEVAL CASTLE

It took me seven years to be built by nine hundred men!! I mean that's a bit long for so many people. And then all winter long I just sit there in the cold, not being worked on at all. Stone and mortar are not the warmest materials, I tell you!

When I was finally finished other things began to happen. These lords moved in with their families, servants and all their junk. Junk-jewels and rugs and furniture that fill up my rooms **soooo** much I can barely breathe!! The horses are always riding over my drawbridge. It is rather annoying to have hooves pound into you every five minutes.

One good thing about being a castle is the jesters. They often make me *CREAK* (a castle way of laughing), even if the king doesn't laugh. Rover was a funny little dog too. He danced around in a hat and shoes. He was **soooo** cute!! Sometimes I think that the king just cares about money and battles. He never seemed to have any fun. I mean, my king, the king who lived in me, never laughed. He never even laughed at Rover.

Battle is the bad thing about being a castle. Enemies are always attacking and shooting at me. Boiling water is dropped off me onto the enemy. All the dead bodies hang around. The problem is, I hate violence. I just can't stand it, and most people laugh but it is not funny. When the Queen says to the King he has a heart of stone it offends me a lot. ***I have a heart of stone,*** but I am not brutal the way he is!!!

Recently people have fixed me up and made me into a tourist attraction. It is really quite flattering. I used to be a place of battle and (dumb) grand meals, but now people come just to admire me and learn my (supposedly) interesting history.

I am now around 700 years old. I just wish little Rover was still around. People think that castles are great but sometimes I wish I were a dog instead.

To Whom It May Concern (in other words: **EVERYONE**),

Child labor is used by even the most popular companies. Nike and Wal-Mart both use children to make the clothes you might be wearing. Child labor is when companies or people force children to work long hours without enough breaks. If kids pee in their pants, they are humiliated and have to wear diapers, no matter what age. The children's wages are very low, only $47.00 a **month,** plus about $12.00 deducted for health care. However, even though the children **have** health care, they are not permitted to go the clinic for check-ups.

So you might be thinking, why don't they just quit? These children come from poor countries. In some cases, their families are starving and they need the money so that they and their family can live. Child labor is illegal in the United States, but it can be practiced in some other countries. Poor countries, with no other jobs to offer. Child labor is bad, but only because people make it. If the companies gave the kids regular breaks, a time for school lessons, and reasonable wages, then child labor wouldn't be quite as bad. But, no. Fourteen year old girls are made to work until midnight 7 days a week. Girls have to pay for two pregnancy tests if they suspect to be pregnant. If the test is positive they are fired. One seven year old girl worked in a sugar cane fields for two years. She dropped out of, school because her family couldn't afford it. She says her father is already dead, and that the money she earns isn't enough for food.

To make these children's lives better would be to the expense of the "masters" of the companies. Oh Yes, it would be **ABSOLUTELY HORRIBLE** to give up some precious **MONEY** so that a child's hard life could be a bit better. (These words should be read **DRIPPING** with sarcasm). But, really, the owners of Wal-Mart and Nike have to be heartless to have this happen.

Think about what you have just read. I am not saying "You have to **STOP** buying their clothes - **OR ELSE** . I am just saying that something can be done. You should write a letter to Nike or Wal-Mart if you wanted, or you can visit these web sites for more information:

http://www.Saigon.com/~nike/reports/report7.html
http://www.hsph.Harvard.edu/gallery/intro.html

http://cgi.pathfinder.com/time/magazine/1998/dom/980330/business.taking a look 7II.html

http://www.nlcnet.org/rtk/rtk2.htm
http://www/unicefusa.org/issuesa7/jan97/labor.html
http://www.freethechildren.org
http://www.digitalrag.com/igbal/index.html

Classroom News

Steve Maleski, who works at the Fairbanks Museum in St. Jay is known mainly as the weatherman, but he actually has multiple jobs. Besides forecasting weather, Steve teaches classes to school kids. He taught my class a lot about kites. He showed us slides of all different kinds of kites in history, how a kite flies, and then we got to make our own kites. Steve also helps run the nature corner during the summer.

As a weather forecaster, Steve uses basic tools: weather instruments (anemometer, thermometer, barometer), a computer, pens, pencils, an eraser, and paper. Steve gets his information from a computer. The information he gets is available to any forecaster; it comes from a main weather center.

To prepare for broadcasting over the radio, Steve says, "I brief myself." He familiarizes himself with what has happened in the last 24 hours, and why. He checks the computers to see what they are saying about the weather. Then Steve compares the information with what is actually happening, making sure that the computers haven't gone off on their own "dreamland."

Weather changes from day to day. Even skilled weathermen can not always be sure what is coming next. They are always on the lookout, though. Watching and waiting for weather.

Ropes Course

The Ropes Course at Lyndon State College was about ropes, but it was more about something else. Teamwork. Everything we did required more than one person. It required the cooperation of everyone. Most things we did appeared very hard, but as we worked together, everything became easier. With everyone's ideas listened to, we had a better sense of what to do as a team.

One of the activities we did was called the Spider Web. It was a big web about as tall as I am (around 5 feet). It was made up of lots of holes, all different shapes and sizes. Most of them could <u>just</u> squeeze somebody through. The object, of the game was to get everybody through the web without anyone touching the rope, <u>but</u> once a hole had been used, we couldn't use it again. There were only 5 holes on the bottom, so we knew we had to levitate people through the top. Everyone had different ideas. We had to figure out who should go first, how they should go, and which holes would fit them best. The most fun of it was when I got levitated through the whole in the web. All the people

handing me through, and the others pulling me towards them. It was also fun to finish it, even though we couldn't play anymore.

I think the day at the ropes course really pushed us into teamwork. We did lots of different things, but all of them had some teamwork hidden underneath the fun. In one day we worked together in tons of different ways. We went to the ropes course thinking we would learn about ropes, but we learned more about teamwork. The two things I would like to do differently next time are be able to rest more and to do the high ropes. But even without those things, the Lyndon State College Ropes course was pretty good.

How to Feed and Water Chickens

For those of you who want to know how to feed and water chickens, here is a step by step procedure. But first, Materials:

- Chicken feed and bin
- Scoop
- Water tubs
- Water (pump, hose, etc)
- Buckets
- Strength (yours)

Steps for feeding:

Step 1: Pick up scoop and go over to feed bin.
Step 2: Open top of bin and put down top.
Step 3: Thrust scoop into chicken feed and bring it up in an arch so that food comes up in the scoop.
Step 4: Walk over to feeding space with food and sprinkle food all over the ground.
Step 5: Go back over to bin.
Step 6: Repeat step 3 and 4.
Step 7: Repeat step 5.
Step 8: Close bin and return scoop to original place.

* Note: depending on how many chickens you have, you may want to increase or decrease the amount of food that you sprinkle on the ground. You may want to do the same thing with the water.

Water steps:

Step 1: Take a bucket and go over to pump (or hose).
Step 2: Hang bucket on pump (or hold it underneath).
Step 3: Turn on water.
Step 4: Wait until bucket is full.
Step 5: Turn off water.
Step 6: Remove bucket.
Step 7: Go over (with bucket) to water tub and pour water in. *
Step 8: Repeat (optional)

*(How to pour: hold handle of bucket in one hand and bottom in the other and tilt so water comes spilling out.)

A Proud Taste for Scarlet and Miniver

By E. L. Konigsburg

Eleanor of Aquitaine was a strong women who did as she wanted. She had a very full life, filled with splendor, gaiety, and sorrow. Eleanor's book, A PROUD TASTE FOR SCARLET AND MINIVER by E. L. Konigsburg, is filled to the brim with interesting characters, and all sorts of events. From a Crusade to a dungeon, Eleanor of Aquitaine, Queen of both England and France, did it all.

The story begins in 20th century heaven, where Eleanor is waiting to see, after all these centuries, whether her husband Henry will be allowed Up from Hell. While she is waiting, she is joined by three friends. Somehow Eleanor's adventures during the Middle Ages come up in the conversation and each is obliged to tell a part of her life.

In the beginning of Eleanor's story Abbot Sugo tells about the Crusade that Eleanor and her first husband Louis (King of Franks) went on. After leaving the beautiful city of Constantinople, the crusaders traveled alone thróugh the mountains. For spending the night, Louis chose a barren spot high up. Eleanor spotted a much greener valley, and ordered everyone to descend while Louis was away. No one noticed the enemy up in the surrounding mountains. Everyone was hot and tired so they removed their armor. When they reached the bottom, the Turks attacked. Most of the French army was injured or killed and Louis was FURIOUS with Eleanor.

Although this book was not very exiting (as you can see above), the characters were interesting and real in this Medieval book A Proud Taste for Scarlet and Miniver.

The Rural School and Community Trust (Rural Trust) is a nonprofit educational organization dedicated to enlarging student learning and improving community life by strengthening relationships between rural schools and communities and engaging students in community-based public work. Through advocacy, research, and publications, the Rural Trust strives to create a more favorable policy environment for rural community perspectives on schooling, for student work with a public audience and use, and for more active community participation in schooling. Founded as the Annenberg Rural Challenge in 1995, the Rural Trust today partners with place-based education programs in more than 700 rural elementary and secondary schools in 33 states.

The Rural School and Community Trust
808 17th Street, NW, Suite 220
Washington, DC 20006
Phone: 202-955-7177
Fax: 202-955-7179
Web: www.ruraledu.org

The Rural School and Community Trust
Research and Evaluation Office
14 Story Street, 2nd floor .
Cambridge, Massachusetts 02138
Phone: 617-495-5191
Fax: 617-495-5908